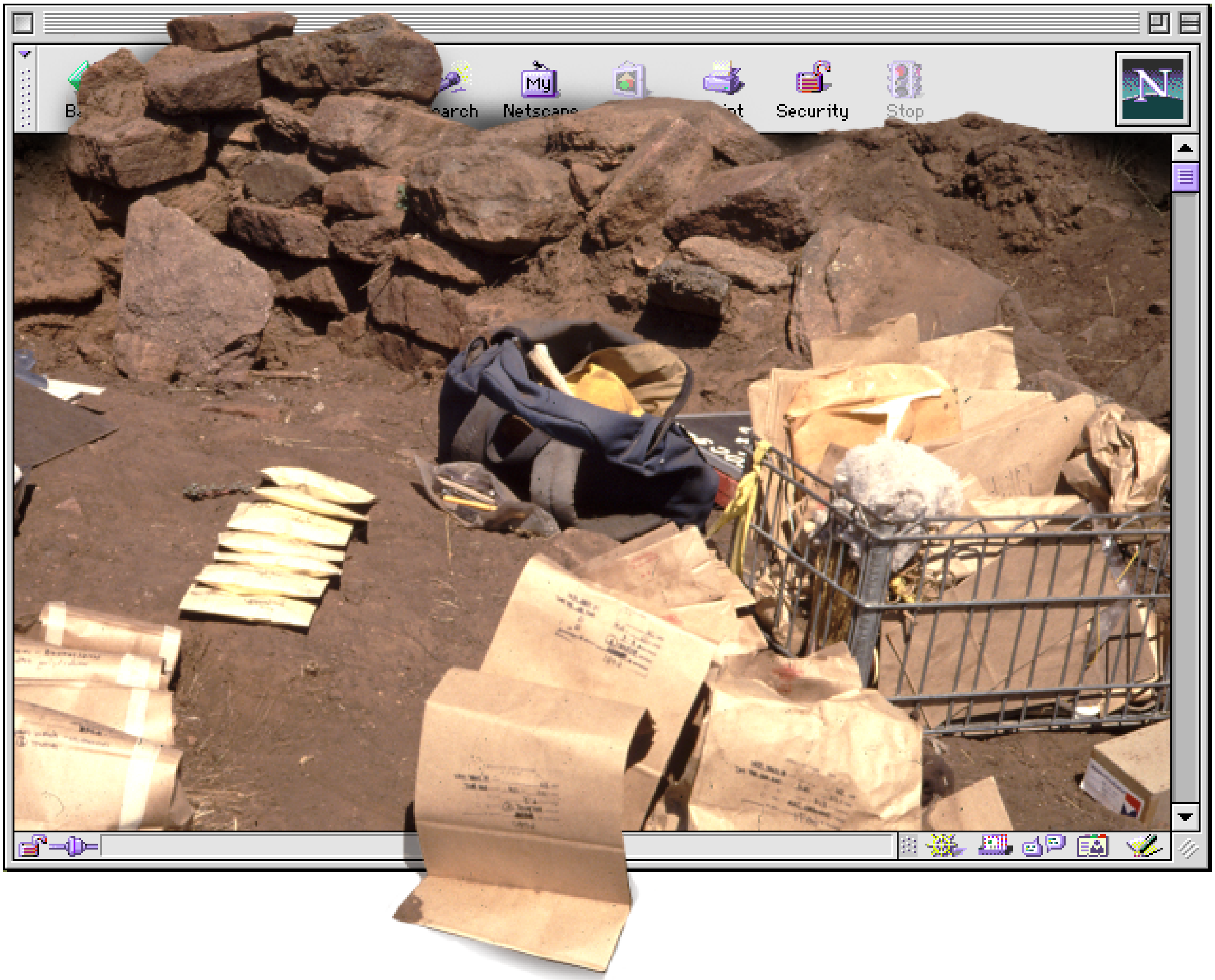
Delivering Archeological Information Electronically

Papers from a symposium presented at the 64th annual meeting of the Society for American Archaeology



National Center for Preservation Technology and Training

National Center for Preservation
Technology and Training

National Park Service

U.S. Department of the Interior

Delivering Archeological Information Electronically

Papers from a symposium presented at the 64th annual meeting of the Society for American Archaeology

March 25, 1999

Co-sponsored by

National Center for Preservation
Technology and Training

Publications Committee of the
Society for American Archaeology

Archaeological Data Archive Project

Editor
Mary S. Carroll

Telephone
318/357-6464

Facsimile
318/357-6421

Electronic mail
ncptt@ncptt.nps.gov

World Wide Web
www.ncptt.nps.gov

PTTPublication
Number 2000-02

Preface

MARY S. CARROLL

Promoting electronic access to preservation information is a major focus of the National Center for Preservation Technology and Training's mission. This focus is embodied in Goal #3 of NCPTT's Five-year Plan — "Increased and improved access to preservation and conservation information and user knowledge of electronic means to obtain information."

Opportunities for providing electronic access to archaeological information have broadened dramatically in recent years. Electronic mechanisms enable quicker access to — and more illustrations in — the kinds of materials traditionally published on paper. In addition, the potential for publishing materials that generally have not been published at all, such as data sets and field records, is greatly expanded. This publication began as a symposium at the 64th annual meeting of the Society for American Archaeology in Chicago. The papers included discuss a variety of information types and access systems in order to assess the utility of various electronic means for the dissemination and utilization of data that are important to the archeological community.

Jointly sponsored by the National Center for Preservation Technology and Training, the Publications Committee of the Society for American Archaeology and the Archaeological Data Archive Project, the symposium was entitled "Delivering Archeological Information Electronically." The session — co-chaired by Mary S. Carroll and Harrison Eiteljorg, II — included nine presentations and two discussants. One presentation from the session — Richard Leventhal's — is not included in this volume. Eight presentations from the symposium are published here in the same order as presented at the meeting with one exception: The comments of the two discussants — Harrison Eiteljorg, II and Mark Aldenderfer — are published as the introductory and closing articles.

As archeology and society as a whole rely more and more on digital information, issues and implications arise — many of which are discussed in this volume.

- How do we ascribe authorship to complex electronic works? And how do we encourage academia to view electronic works in the same light as print publications?
- Accustomed as we are to citing print publications, how do we cite electronic sources, especially constantly changing Web sites? Though mundane, the troublesome nature of this question became clear to me as I put this publication together!
- How do we best plan and design Web sites so that users can access information easily and efficiently?
- How can we use electronic media for purposes other than research or information dissemination — such as teaching and communication?
- What direction will paper journals take as electronic journals develop? Will they disappear altogether or will they co-exist with their digital counterparts?
- How do we preserve the digital information that we are creating? Along with the focus on electronic dissemination of information comes a responsibility to preserve digital data. Preservation of digital data is a complex and critical issue that involves both the permanence of the media on which the data are stored and the rapid changes to the technology used to access the information.
- How do we ensure that the digital data are not only preserved but also accessible, usable and understandable?
- How do we ensure access is provided to the appropriate audience? How do we protect data of a sensitive nature? When is it appropriate to open access to all and when should access be restricted?
- Is the technology used appropriate to the task at hand or is it a case of being swept along in the race to use state-of-the-art systems? Are we engaging in overkill?

The goal now is to continue the discussions and to include industry and the wider preservation community. These issues will not resolve themselves but require systematic, long-term attention so that archeological data — and any digital data — will not become lost remnants of the past.

Acknowledgments

Many thanks go to the participants of the session for their insights and enthusiasm, to the Publications Committee of SAA and the Archaeological Data Archive Project for agreeing to co-sponsor the symposium, and to my co-chair for his assistance in soliciting speakers and focusing the session.

Introduction

HARRISON EITELJORG, II

The title of this session, "Delivering Archeological Information Electronically," suggests to me two rather different issues — electronic delivery, and electronic or digital archeological information. I would like to use that dichotomy to discuss what we have heard, and not heard, in this session.

Electronic Delivery

First electronic delivery, the topic that received the lion's share of our attention — appropriately enough, given the way the session was organized. It seems to me that we have had a superb discussion of the state of the art today, from the very practical approaches of Mary Carroll and David Carlson to the more theoretical approach of Christopher Chippindale.

Christopher Chippindale was the only one who mentioned the possible desirability of hybrid publications — partly on paper and partly in electronic form. I am inclined to agree with him that such publications will be more common than we have thought.

Donald Sanders and Richard Leventhal talked about the process of taking an archeological work and transforming it into something unique to the electronic world, complete with complex and wondrous virtual reality effects. Donald Sanders spoke at some length about the questions of authorship and authority raised by the process, and it seemed to me that he showed well how the problems could be dealt with — not overcome — since the electronic world must involve different realities, but dealt with effectively.

Richard Leventhal showed how many differing data types can be combined in electronic resources to produce a unique document. He pointed out how many different skills — and the people possessing those skills — would be required to complete an electronic publication.

Jim Farley and Peter McCartney showed us a different kind of electronic delivery — in their cases less publication than direct computer access to primary data. In Jim Farley's case, the data are GIS data sets, and the point is to provide a system that will make direct access to the data themselves possible, with little mediation, for the wider public. As a result, some issues of standardization are very important, and those issues figured in Jim Farley's presentation. Fortunately for that project, GIS data are very important in the commercial marketplace. As a result manufacturers of GIS programs are being — and will continue to be — driven toward standards for commercial reasons that do not apply for very many other data types. Producers of database management systems and CAD programs do not have equivalent commercial needs to cooperate in seeking agreement on standard data access systems. Nonetheless, the planned access to GIS data provides a good example of the kinds of standards many of us would hope to have, not only for GIS but for many other data types as well.

Peter McCartney (and I when representing the Archaeological Data Archive Project) cannot now provide the kind of access that Jim Farley is preparing for GIS. That may come but, in the meantime, Peter McCartney indicated his concern to preserve primary data so that access can be provided — and possibly easier access like that envisioned by Jim Farley in years to come. Peter McCartney also stressed the importance of metadata, the data that will make indexing possible and, consequently, permit us to find the information we need on the Internet. Metadata also will help future users of the data files by providing authentication, information about file types, migration history and so on.

Terry Childs discussed the work of the Federal government in presenting digital information. She and David Carlson were the only ones to discuss listservs, and that brings up the unstated assumption in this session — that electronic delivery equals the Web. Given the experience we have had thus far with computers and the Internet, I think that it is very dangerous to make that simple equation of electronic delivery and the Web. If there is one thing of which we can be sure, it is that there is no way to predict the nature of the computer or the network very far into the future. Therefore, I think we must watch ourselves and try to think more broadly about the kinds of electronic delivery that may be in use in the future.

Terry Childs emphasized the need to consider the audience when designing information presentation systems. The government has many potential audiences; it must deal with all those constituencies, and the experience gained there can be useful to anyone offering information over the Internet.

David Carlson had a very specific audience in mind — students in undergraduate classes. David's practical suggestions were valuable and helped to remind us that one must very carefully tailor the message to the audience.

Hugh Jarvis, speaking about the use of electronic publication for journals, made the strongest statements about the coming dominance of electronic dissemination of archeological data. While I am not sure I would agree entirely, the basic point is incontestable. Whether we prepare or not, we will soon see enormous increases in the quantity of material presented electronically.

I have saved my colleague and co-chair of the session, Mary Carroll, to last. Mary Carroll reminds us that we must not only plan carefully but also be clear about our goals. Then, goals in mind, we must measure our progress. We dare not march blindly forward without examining the results of our work. This may be the most important message of all. As important — and enjoyable — as it is to work in these new and fascinating areas, the point is to accomplish specific aims. We need to be sure we understand our aims and can measure our achievements. Then we must honestly do the measuring.

Digital Archeological Information

Now let me turn to the second of our issues — electronic or digital archeological information. It seems to me that we have given this area less attention — less than deserved and required. What kind of information is appropriate for what purpose? In particular, I am concerned that we may be tempted to revel in the possibilities of new and impressive "gee-whiz" technology, whether it is useful and economically justifiable or not. I am also concerned that even appropriate, less exotic technologies may be used in ways that are as potentially harmful as helpful.

I would like to approach this issue from two angles. First, what level of technological wizardry is appropriate for what purpose? Second, when we let computers work for us, are we aware of the potential problems that may accompany the benefits?

We start with the question of appropriate technology, and I begin with a story. Yesterday I went to a west Chicago public school where my daughter teaches. I took slides of an excavation so that I could talk to fourth- and sixth-grade students about archeology and a particular excavation. The promised slide projector turned out to be a non-functioning filmstrip projector instead. I was obliged to talk for better than an hour to each of two groups with only a blackboard as a visual aid. Despite the absence of slides, I was able to keep the children's attention. Now none of us would choose to make a slideless presentation of an excavation, but my point is that the tools we expect may not be as necessary as we think.

Put that in an electronic context. If we are presenting information, do we need all the bells and whistles? For instance, does a scholar need a virtual reality presentation or only simple tabular presentations of data — or both? Does a student need the tabular data or the virtual reality presentation — or maybe just good renderings of long-gone structures — or all of those presentation types? If we are showing the results of a survey, do we need colorful maps and charts or something much simpler? I suspect each of us can come up with an example of computer overkill. Given the costs — something almost all of our speakers mentioned — this is an important question.

With each sophisticated technology there is also the question of the appropriate extent of its use. Which bells and whistles are needed at any given moment? While it may be useful to show a house as a virtual world in which a viewer may move, at least metaphorically, the same person viewing the same excavation may not need to see every pot (or any pot) with that same level of realism or to have other objects from the excavation available as virtual objects to be lifted, moved, and rotated. Virtual houses or pots are only stand-ins for the real things; after all, how close to the real items must the stand-ins be? The issue, once again, is partly cost. When does a virtual reality pot serve a purpose so necessary that the expense of its creation is justified?

The use of extremely realistic presentation systems brings an added problem: how clear is the distinction between the real item and its stand-in if the representation is photographic in its realism? Of course, that distinction becomes more and more important as the proportion of restored material increases and the proportion of real, extant remains decreases. Highly realistic presentations of mostly restored artifacts, buildings, or sites can be hard to resist — even if they are not supported by the evidence.

There may also be problems with the use of well- and appropriately presented data. Imagine CAD models, databases, or GIS files as parts of electronic publications. In order to use the data effectively, a user will need some skill and experience with the appropriate software — CAD, database management systems, and GIS programs. Yet how many of those here in this session concerned with technology have experience with all of those program types? Too few professionals — and fewer graduate students — have enough familiarity with different software types to access effectively the many different types of digital data that are available. If that is true, what should we publish or make available in digital repositories and when should we begin to do so? Should we wait for the technological sophistication of users to catch up with the technology, or should we make available extremely sophisticated types of computer data on the assumption that the users will ultimately learn enough to be able to use the data we store today?

As Mary Carroll pointed out, we need to specify our aims and to measure our progress. That is as true of the kinds of data presented as of the forms of presentation.

Now let us turn to the second of my concerns about digital data. I believe there is a problem — or a potential problem — with the way we may obtain and use digital data from computerized repositories. Peter McCartney and Christopher Chippindale talked about primary data and capta — I take the terms to be equivalent (and I like the term capta). The next level of information is those capta with context — something Peter McCartney called information. Information in a wider context may become knowledge, and knowledge may then be summarized. What is probably obvious to all is that the summary from one level of study becomes the capta for the next level, and the summary from that level becomes the capta for the next. The cycle will repeat itself endlessly, and at each step — from capta to information to knowledge to summary to new capta and so on — there are rules, processes and assumptions that guide the transformation. Those rules, processes and assumptions are crucial to the whole enterprise; they determine the outcome of data searches, aggregations and summaries. The potential problem lies in the fact that the nature of the transformation may too often be ignored in the haste to obtain results. However, the nature of any transformation must always be clear, explicit and transparent. That is, the scholar must always be able to know how information was generated, starting with the capta at the very beginning of the process, and including all the transformational processes thereafter — every rule, process and assumption that participates in the data transformation. Most scholarly publication, after all, is an attempt to show how explicit transformational processes allow certain information to be transformed and made meaningful in a new context; the nature of the transformation is often the point at issue.

I worry that computers make it far too easy to assume that the transformational processes are computer-generated and, consequently, above reproach. We cannot permit that. Each and every process in this transformation chain must remain not only explicit but also transparent, and we must train ourselves to pay as much attention to the processes that generate the information as to the information itself. Otherwise, we risk having and using information that has been created by processes that are unacceptable to us or based on rules or assumptions we do not accept - perhaps stemming from old and outmoded models, possibly using inappropriate statistics, maybe involving data aggregation assumptions we believe to be faulty.

Digital data, then, must be treated with great care. The chain of creation must remain with the data, and potential user(s) must always be aware of the way the chain was forged. One bad link will render the remainder useless. Those of us who provide digital data must take great care to be sure that we are able to provide more than the data. We must be able to provide an "audit trail" that connects the first capta to the latest summary so that the archeological information is not only digital, it is also understandable, reliable and useful.

Conclusion

Like archeological information on paper, the archeological information we get in digital form must be both appropriate and trustworthy. The forms we can obtain must be the right ones for the job — cost effective and usable. At the same time, what we receive — capta, information, knowledge or summary — must be dependable, resulting from transformations that we can understand and evaluate.

In sum, we have seen here many interesting perspectives on electronic delivery of archeological information. It is clear that we will be using electronic forms of delivery more and more in the future; it is clear that we can do many new and exciting things in the process; it is clear that new paradigms must be developed. At the same time, it is equally clear that issues of cost remain to be determined in many areas, that we must resist the temptation to use the technology for its own sake, and that we must consider the skills required of users. Finally, we need to be more aware of the potential problems of providing data that have been manipulated by computers for us — but not in ways permitting and requiring our inspection and examination. In each of these areas the key to moving successfully forward may be found first in having real aims and goals, second in planning appropriate measurements of our results, and third in carrying out the measurements. Some of these issues, however, require more open discussion among interested scholars — particularly the issues surrounding the question of appropriate digital forms for archeological data. Perhaps these issues would be appropriate subjects for another SAA session.

The Nature of Data in Paper and in Electronic Media

CHRISTOPHER CHIPPINDALE

Data is not a fixed given but is itself created and shaped to make knowledge. The forms that transmit knowledge also create and shape it. Now that electronic media are established alongside paper, we can see how each form gives a certain character to archeological knowledge both inside a specialist community and out to a wider public. So we can begin to plan for the mixed-media future in which the Society for American Archaeology's own publishing will have a role.

If the book is the standard or "pure" printed medium, then the Internet is at the other pole — the electronic medium least like print. The CD is an intermediate form, sharing features with both. A CD resembles a book in that it is centrally manufactured, physically distributed and purchased by the user. Yet, like the Internet, it is electronic and potentially interactive. So I address here the two poles — the printed book and the World Wide Web — rather than the CD medium in between.

Skeuomorphs: From Old Forms to New

Archeologists know about skeuomorphs — the tendency for objects made of one physical material to mimic the forms fitting to another — from their own specialized studies. One context for skeuomorphs is the introduction of a new material or a new technology, which characteristically mimics a pre-existing form until such time as its own character emerges. So it was that early ceramics often follow the baggy shapes of leather or basketry containers and early railroad cars repeat the designs of the stagecoaches.

Electronic publishing is repeating this tendency. The Web unit is called the "page" — although the technological cause for print to be organized in pages is the medium of folded paper sheets with constraints which do not apply to electronics. PC operating systems mimic a paper office, with a "desktop" and "documents" stored in "files" grouped together within "folders." These helpful paper and print analogies are damaging since they fail to match the different structures of the electronic world.[1] The varied material on a CD or a Web page has no simple linear order as a book does, but more likely follows the pattern of a branching tree. The convention is emerging of either a table of contents on the model of a book — or a fuller set of conventional front matter[2] — or of a site-map to show what is there. Against this is the ideal of a Web site, Ruth Tringham's *Chimera Web*[3] — which does not so much guide you in a specific direction as try to help you explore an unknown.

Novel[4] electronic forms are now emerging. One is the "webcam," the online camera reporting activity from some remote place. Matthew Spriggs attempted an archeological example in 1998[5] to report an excavation on a remote Pacific island each day to his students back home at the Australian National University.

Skeuomorphs resolve themselves as it becomes clear what is more genuinely novel — as a webcam appears to be — and what closely follows an existing form. One sees this in the electronic journals like *Electronic Antiquity* and *Internet Archaeology*[6] as they decide how much to create an electronic simulacrum of a paper journal, differing only in its means of transmission, and how much to change the form. Novelty has its risk. *Internet Archaeology* emphasizes its commitment to the virtues of print journals alongside its wish to develop novel elements unique to electronic formats. One can easily and repeatedly update an electronic publication or append readers' responses in a way that cannot be matched by revised editions of a paper book or article. At the same time, some old issues are fundamental to publishing — and to knowledge! — such as the question of quality control; these will endure.

An instance in which an old form is adapted to a new medium is the moderated discussion forum — such as AegeaNet,[7] which resembles the standard newspaper or magazine formula of "letters to the editor" — where new issues are raised or old ones revisited in a supervised forum with a controller who decides what will and what will not be accepted for publication. An instance of a decisive difference between old and new forms is the perception that paper publishing is expensive and electronic publishing is cheap. This belief is related to the fact that print publishing is usually done by a third party, while much electronic work is self-published by authors.

Costs: Actual and Perceived

Although the costs of carrying out some academic research, and then of publishing it, can be calculated with some precision, there is enormous variation in how these costs are met and how they are therefore seen by the ultimate customer — the individual who acquires knowledge.

First, what matters is more the *perception* of cost than the reality. The key decider becomes bookkeeping rules of who pays for what. Browsing a colleague's office shelves recently, I was struck by how few of the books were new. He explained, "I don't pay for books any more — I just photocopy what I need." As he experiences it, photocopied materials — or materials printed from a Web file — are free whereas books are expensive. A publisher's editor, knowing the fixed costs that have to be recovered by selling sufficient copies, will scarcely agree.

Photocopying is a simple instance of deceptive perceived costs. A university teacher can buy a book for a certain number of dollars or request the library to buy it; that price consolidates the varied expenses involved in making the book. Once the book is by some means to hand, it can be reproduced "free"[8] by the professor because the many costs involved are scattered between several places. Costs incurred by the university department are hidden in other bills — capital cost of the photocopier, maintenance, toner, paper — which are treated as overheads. If the professor himself stands over the machine to photocopy it, there is another cost, that of his time; at $40 an hour and more, that is not a trivial amount in relation to the price of the book. Against that example, however, is the experience of those university presses like University of Arizona, which have tried putting whole books up on Web pages. This permits a would-be reader to print out the entire volume with — if in a typically indulgent university department — no perceived cost. But Arizona finds that the availability helps rather than reduces sales of the printed book.[9] Readers, it seems, are enticed by what they see on screen, or by the fragments they print out, to the point that more of them purchase the book.

Second, the conventional charging of academic work and of academic publications is itself a partial account of the actual costs involved. Consider a substantial field research project, one having the general character of the many campaigns conducted in the eastern Sahara by Fred Wendorf and colleagues, or a major single excavation such as Franchthi Cave, Greece — published largely as a collected set of "fascicles," each a volume on one aspect of the excavation and the material.

The knowledge generated by projects of this nature is made available as printed books to which electronic publication on the Web is a potential alternative. How do the two media compare?

The major cost is the research itself and within that research cost the major element will be labor. But most of that labor cost is not charged to the project. It is instead supplied by the universities paying the salaries of the professor who carries out the research and by individual students and associates who work without charge — or, by their course fees, actually pay to work — or who work for much less than market rates. Universities, in providing laboratories, also pay much of the immediate overhead costs of facilities — heat, light, and so on; individuals do the same when they work from home. Universities also provide the infrastructure of libraries, reference materials and so on. Research funding, from a body such as the National Science Foundation or the National Endowment for the Humanities, only pays that — often small — portion of research costs that arises as direct and specific bills — travel, accommodation, locally hired labor, working materials, shipping, the specialized advice of expert analysts and so on. Even the element of "overhead" does not often carry the full cost.

Conventionally, this framework changes when one moves on from the research itself to publication of the research results. "Gray-literature" reports often are available without fee from the publishing agency, which bears all the bills. With a commercial or quasi-commercial press, the "full costs" of the publication — editorial, overhead, sales and marketing, origination, paper, printing, binding, warehousing and distribution — are to be recovered by sales. For a non-commercial press, such as a museum monograph series, most full costs are provided for and a smaller view is taken of the allocated costs, which may be the bare manufacturing bills from the printer. Since the publications are very specialized, not many are printed and even fewer sold, so the price is often very high. Yet even high book prices do not pay for the research itself. By simple rules of supply and demand, expensive books on specialized subjects sell very few copies. The benefit of all that research work is decisively weakened at this last stage in the process, because the book is in restricted circulation. One would do better to make a nominal charge for the book and see its publication as an integrated element to the cycle of making and circulating knowledge — a cost like any other.

Before all the oddities and exceptions, the fundamental cost conventions of paper publication are these: the large research costs are not covered by book purchasers but carried in some other way by the producer; some or all of the immediate publishing costs are carried by the consumer.

Famously, the costing and business economics of the Internet are strange. Central elements, such as the software of Web browsers and e-mail components, are produced by commercial companies — yet often no charge is made for them.[10] User access by Internet service providers is usually charged but may be free. A university user will usually perceive Internet access as free because the bills are covered by the institution or department, not charged to the individual.

Let us look again at the same publication of a large and specialized research report, if it instead takes the form of a Web page created in-house by the research team in a university department and hosted on its server.

Much of the work is the same as for print publication, but who does it and how it is charged will vary. Peer review, revision in light of review and publisher's opinion and advice, and copy-editing are in theory the same procedures in each case. In practice, the in-house publication is more likely to be published as written and received, with less effort being put into review and revision. Typography, design and page mark-up in print publications — invariably done by electronic methods today — have a close equivalent in electronic page design and mark-up. The work in the print version will be done by publisher and printer, and so charged, but in the electronic medium design and mark-up are done by the author — who makes no charge — or by students or assistants who are paid not at all or at less than commercial rates. Origination and proofing is cheaper for electronic media, and again the labor may not be charged. The equivalent of physically manufacturing the book is its posting on a Web page, which carries a cost — but the cost probably will not be charged or even perceived. The equivalent to distributing the physical book — moving it from bindery to customer — is dialing-up the Web site's URL. Access is often cheaper and perceived as free and is a cost borne by the reader rather than the publisher.

Books require no equipment of the reader. Electronic publications require the reader to have equipment — CD-equipped PC, modem and telephone line, software — not usually thought of as part of the publication costs.

Even before the oddities of the Internet as a business environment are taken into account, one can notice what happens when *distributed* reproduction takes the place of *centralized* reproduction — that is, when individual users in scattered places photocopy or print material rather than receive a printed book from a central publisher. Laser output is expensive — as the work is printed one sheet at a time — by comparison with the efficiency of centralized book printing. Distribution costs are sufficiently low, given an efficient postal service, that central reproduction is cheaper overall. Although various schemes exist where a charge is made for photocopying copyright material, most photocopying is done without such payment being made, so the central indirect costs are borne by a reduced number of — in the case of a journal — paid subscriptions.

A few words are useful here about costs for CD publication. Compilation and layout of words, images and other material is comparable to creating a book to the point when it is ready to be printed. Overall fixed costs would be lower for a CD if it contained the same quantity of matter as a book. But a CD's larger capacity – 650 MB, about 6500 text pages or 1000 color photographs – means the "typical CD" can be fatter in its content than a book – which is the attraction of the format. Much more material at rather lower unit cost means an overall greater cost.

Making the master CD and duplicating copies appears cheaper than the equivalent work of plate-making and printing, noticeably at the print runs of not many hundred which are the universal rule for specialized archeological publications.

Who pays? The University of North Carolina Press paid the basic costs of pressing CDs from master files supplied in final form by the authors. This is a decisive shift of costs from the publisher to the author.

What are the decisive points here?

- There is a change in cost structure in which electronic is perceived as cheaper.
- There is a reduction in the role of the publisher as the third party intermediate between author and reader.
- There is some shift in costs from the reader towards the author and the author does more.
- The fixed costs — those involved in creating the "master" original — are higher if the electronic potential is taken up, while the variable costs of making each copy available to customers appear to be reduced.

Quality Control

Quality control will always be in relation to some perceived framework of knowledge. Some studies of the past use frameworks decidedly other than those of the academic community of research archeologists. On the Thames & Hudson list, alongside the high quality archeology books, are varied books of a mystical, "New Age" or "modern Celtic" spirit, often containing little or even false knowledge.

In the accepted framework of print publication, the author's work is assessed by the publisher as an independent party; whether it is published and in what form follows from that judgment. Academic papers are assessed by peer review, academic books by some combination of peer review and opinion of the proposal's commercial potential. Peer review is uneven and necessarily subjective. A reviewer may be supportive or skeptical of the merit of the research, broader- or narrower-minded in their view of what is good work, tolerant or intolerant of slips and weaknesses in detail and in presentation, or inclined to be generous or not when it comes to a proposal with middling merit. Editors and publishers are equally varied in their view. Many established presses, especially university presses, maintain a habit of demanding copy-editing. Much peer review is weak or even nominal, in that it is done by colleagues of the author who know well the work under review and its author, who are themselves researchers of the same attitude and orientation, and who approach it expecting it to be good. The set of publications in archeology which have been rigorously peer reviewed by colleagues at a distance from the work is not the same as the set of publications in archeology found good and useful. John Maddox, whose many years at *Nature* make him the premier journal editor of our age, reports that after a while he stopped refereeing papers offered to *Nature* by the astronomer Fred Hoyle, a highly original thinker whose ideas were not well-received by colleagues in peer review. Maddox printed them anyway and does not regret that.

Although the refereeing process is intended to control quality, it also reduces quantity. Rejected papers and books are often abandoned rather than published — with or without revision — in another place.

Alongside peer review is editorial evaluation, what can properly be called "news values." As I write, on 27 January 1999, the day's national newspapers report the events of high importance — the next step in the Presidential impeachment, Pope John-Paul's visit to St. Louis, the continuing fears for the impact of "Y2K" on computer systems. Other events of yesterday, from the minor collision on Temple Street in Salt Lake City to the present author's research seminar offered to colleagues at the University of Utah, are rightly overlooked. Explicitly or implicitly, academic publications also work by news values.

Consider *Journal of Archaeological Science* and *American Anthropologist*, both journals of established reputation — *Journal of Archaeological Science* published commercially by Academic Press, *American Anthropologist* by the professional non-profit association. Yet *Journal of Archaeological Science* (editor Richard Klein) and *American Anthropologist* (when edited by Dennis and Barbara Tedlock) held to such different paradigms — "news values" in my terms — as to what kind of work was of merit that it is hard to think of a paper published in one which could equally have been published in the other. This is acceptable, even valuable, because the research community that reads the journals knows and understands these values. That community is also aware that the values change, although abrupt shifts confuse. Whatever the merits of any emphasis, a chopping-about of editorial values confuses readers because it perturbs their understanding of where the journal stands in the intellectual landscape of the discipline.

All researchers in a field are aware of the varied reputations, standards and attitudes of the journals and publishers. Learning the shapes of this "publishing landscape" is a key skill the novice researcher acquires. So is learning to notice the clues, large and small, that increase or reduce a reader's confidence in the quality of what they are reading.

So the mechanisms of quality control are variable and inconsistent in print publication, because there is neither objective and absolute value, nor some threshold to divide simply the good from the bad. Peer review is only part of the story. Some dogs of papers and books are published by imprints that should know better! Equally, some papers and books subsequently considered significant have been published in obscure or unexpected places because the mainstream rejected them. Imprints which are too zealous in peer review may find themselves squeezing the originality out of their contributions, so one arrives at a published text which is safe but contributes little new. What remains the case is that the author must either persuade a third party — or parties — of the merit of their work or the author must fund publication themselves, knowing that the absence of third-party independence and the bad reputation of "vanity publications" may poison expectations.

In contrast with print, the Web operates through a routine of self-publication. The analogue of the print publisher might be the service provider who hosts a Web page; but Internet service providers have little concern with what is on their servers beyond a minimal interest in its being legal or decent. So there is no system of third-party control, no distanced judgment of merit and quality, no independent editor to cut and to shape and to resist the author's inclination to write yet more words in yet slacker prose. On the Web, in fact, there is no boundary of the kind that separates and distances the extremes of formal publication and passing gossip. The clues of format and visual presentation that help the reader of the printed media to sense the standing of any one printed work only weakly apply to the Web, where the most accomplished design may be the medium carrying the gossip and the self-indulgent fantasy. A free-speech ethic and habit coupled with the lack of controls results in copyright not being an actual restriction on the re-use of material.

Given these freedoms, I am astonished by how good the stuff on the Web is, how evident is the care taken to get things right and how much good archeological material is available.[11] Ancient Egypt, for centuries such a field for historical speculation and fantasy, is a striking instance where the orthodox archeology shades off into other visions in a way hard to navigate.[12]

A likely way forward will be developing structures within the Internet that will make the standing of the site clear. A print publisher of good reputation will be an indicator of a certain quality when it issues a digital publication — such as Oxford University Press's online journals. But it also will be the case that some personal Web pages will earn good reputations without such external validation — as has been the case in the print. In its time, *I.F. Stone's Weekly* — the little self-published weekly report and commentary on public affairs in the nation's capital — had an authority not to be dismissed. "With a few exceptions every issue of the paper was written, from cover to cover, by I.F. Stone himself."[13]

I introduced this section of my essay with the words "quality control," for that is the term by which these issues are being noticed. I prefer a broader term, such as the "topography of knowledge" to indicate that wider set of issues and judgments within which any measure of quality is set.

Atomization

A distinctive feature of the new topography of electronic media is an *atomization* of knowledge — that is, a willingness to treat knowledge as an accumulation of facts piled together rather than as a synthetic understanding beyond the sum of its component fragments. Much print publication is some kind of accumulation. An academic journal, like a newspaper, is a miscellany of varied reports on varied subjects within some defined field. Electronic media are even more varied and this has consequences.

Behind every factual statement, however straightforward and objective it may seem, are disputable judgments. The population of the United States, my atlas says, is 241,596,000 — to the nearest thousand. Beyond the obvious qualifiers — such as the date of the statistic — there are some less obvious ones. Are US citizens permanently abroad included? Are US citizens temporarily abroad at some defining census date included? Are foreign citizens permanently resident in the US included? Are foreign citizens temporarily in the US included? How many individuals are omitted from counts? The numbers of those of uncertain status are so large, despite a professional census bureau aware of the complications, that the proper counting of the number of people in the United States is a recurrent subject of dispute and litigation.

A central issue — or *the* central issue — in archeological method is that of the "middle range" in formation processes;[14] what is the relationship that links those physical objects that we can observe and measure to the human beings and social entities that an anthropological archeology seeks to study. Every archeological observation comes, or should come, with a sense of fuzziness, whether the list of major settlements of the Maya realm or the counts of different tool types in a stratum at an impeccably excavated French Palaeolithic site. Just one class of archeological observations, the radiocarbon date, routinely is reported with its measure of uncertainty. In truth, uncertainty is attached to every radiocarbon date even beyond what is expressed by its standard deviation and uncertainties surround and cloud most archeological observations, however neutral or objective they appear. Each statement depends on definitions — what is a "flake," a "blade," a "core," a "piece of debitage" in stone-working? — on sampling, on the hazards of taphonomy and on what the observer chooses or chances to observe.

Central to this tendency is the word "data," a short and everyday word whose dangers we overlook. "Data" derives from the Latin "datum," meaning "that which has been given." But data are not given, certainly not in archeology. Rather they have been *captured*, by some effort of studied observation. For this reason I rarely use the word "data" and prefer instead the unused word "capta," which better expresses the truth. We go out in search of facts pertinent to some research interest, and seek to capture them — but we may come back with nothing at all or with observed facts unrelated to our research interest. "Capta" reminds one of that real uncertainty.

Many of the remarks in this essay concern the context of knowledge. Electronic media promise, and already deliver, vastly more facts and factoids — a more contestable and uncertain statement which is nevertheless presented and treated as if an undisputed fact. In principle, the same conditions apply to printed and to electronic media. But both the technology and the emerging habits of electronic knowledge promote a "cut or copy and paste" spirit, in which the context and conditions on which the data depend are speedily lost.

An incentive, noted above, for publishing electronically is that more data can be released. Further, because those data can be copied and manipulated with a few clicks of the mouse, they are more accessible for new study and new interpretation than are printed lists and tabulations, which have to be laboriously re-keyed or scanned. But are those data truly independent of the theses developed in the synthetic portion of the study? How much can they indeed be treated as not subordinate? Here the reciprocal relationship of theory to data comes into play, in which the theory depends on the data, and the data depend on the theory. No empirical rep — certainly no archeological report from the field, the lab or the museum — is a complete or an objective report. Rather, it is a necessarily selective set of observations — those pertinent to the subject of study. Aldenderfer[15] notes a telling case in this respect. Hill's celebrated and influential study of the Broken K Pueblo site[16] was a landmark in developing a self-consciously scientific "New Archaeology." In the modern monograph mode, it presents a mass of supporting data. A re-analysis taking a different approach could — just! — be done[17] by manipulating and re-interpreting those published data. However, in order to compare results properly, Lischka secured from Hill copies of his original data-runs, records more primary than what was published. Dillehay's recent two volumes on Monte Verde[18] — with their many pages, many tables, and many illustrations — present a great quantity of observations and analyses. But their purpose is to demonstrate what was observed and the logical means by which a particular synthesis was made of the Monte Verde evidence. The synthesis depends on what existed at the site, but it also depends on what Dillehay's team chose to recover, record and study and by what means and within what frameworks of ideas.

Knowing that data are not in a simple way either independent of or dependent on theory, the present writer does not welcome the atomizing of knowledge or the increasing removal of the data from the conditions under which they were created. The new Archaeology Data Service[19] now asks British researchers in receipt of funding from the Humanities Research Board to lodge their "data-sets" with the service for other researchers to use. This would ease the tiresome duplication of work that arises when one researcher recapitulates what another researcher has done, but it would also lead to error whenever the context for that data was not properly taken into account.

Acknowledgments

Like others contributing to this SAA session, I have benefited from a January 1999 meeting called by the "Digital Imprint", the project directed by Louie Krasniewicz at UCLA, to develop electronic publishing in archeology. I thank the Digital Imprint team and other colleagues attending.

Notes

1. Nat Tunbridge, "The Human Touch," *New Scientist* 161.2170 (1999): 34–37.

2. R.P. Stephen Davis Jr. et al., eds., *Excavating Occaneechi Town: archaeology of an eighteenth-century Indian village in North Carolina* [CD]. (Chapel Hill, NC: University of North Carolina Press, 1997).

3. <bmrc.berkeley.edu/people/ tringham/chimera.html>.

4. Nothing is wholly novel. This is, seen another way, a live long-distance TV transmission of the kind which has been routine for half a century; it differs in the way the viewer accesses it.

5. Matthew Spriggs. *Pacific Archaeology Teaching Project 1998*. 1998. <artalpha.anu.edu.au/nobarriers>.

6. Mike Heyworth et al., "Internet Archaeology: a Quality Electronic Journal," *Antiquity* 71.274 (1997): 1039–1042, <intarch.ac.uk/antiquity/electronics/ heyworth.html>.

7. John G. Younger, 1997. "Managing 'AegeaNet'," *Antiquity* 71.274: 1052–1054 <intarch.ac.uk/antiquity/electronics/younger.html>.

8. Various rules and schemes exist to discourage photocopying or to make a charge to the user that is returned to the author/publisher of the material copied. The policing rules are not rigorous. Much of the funds generated in payment for photocopying is consumed by administering the charging schemes. Many academic authors, who want their work to be noticed and read, are content for it to be disseminated by photocopies, since the income they receive from licit copies is slight or zero.

9. Chris Szuter, personal communication.

10. So programs like the Netscape browser and Eudora e-mailer are free to many users in their standard form. Their proprietors survive commercially by charging for elaborated versions and ancillary software. Also, Internet software is central to the competition between Microsoft and other software companies who may cross-subsidize from other business.

11. Sara Champion, "Archaeology on the World Wide Web: User's Field-guide," *Antiquity* 71.274 (1997): 1027–1038 <intarch.ac.uk/antiquity/electronics/champion.html>.

12. Lynn Meskell, "Electronic Egypt: The Shape of Archaeological Knowledge on the Net," *Antiquity* 71.274 (1997): 1063–1076 <intarch.ac.uk/antiquity/electronics/meskell.html>.

13. Neil Middleton, *The Best of 'I.F. Stone's Weekly': Pages from a Radical Newspaper* (Harmondsworth: Penguin, 1973).

14. Michael B. Schiffer, *Formation Processes of the Archaeological Record* (Albuquerque: University of New Mexico Press, 1987).

15. Mark Aldenderfer, "The Printed Monograph: History of a Class of Archeological Publication" (Unpublished paper given at the Digital Imprint meeting, UCLA, January 1999).

16. James N. Hill, *Broken K Pueblo: Prehistoric Social Organization in the American Southwest* (Tucson: University of Arizona Press, Anthropological Papers of the University of Arizona, 1970), 18.

17. Joseph J. Lischka, "Broken K Revisited: A Short Discussion of Factor Analysis," *American Antiquity* 40 (1975): 220-227.

18. Tom D. Dillehay, *Monte Verde: A Late Pleistocene Settlement in Chile 1: Paleo-environment and Site Context* (Washington, DC: Smithsonian Institution Press, 1989).
Tom D. Dillehay, *Monte Verde: A Late Pleistocene Settlement in Chile 2: The Archaeological Context and Interpretation* (Washington, DC: Smithsonian Institution Press, 1997).

19. Julian D. Richards, "Preservation and Re-use of Digital Data: The Role of the Archaeology Data Service," *Antiquity* 71.274 (1997): 1057–1059 <intarch.ac.uk/antiquity/electronics/richards.html>.

Author! Author?

DONALD H. SANDERS

In an increasingly electronic world, archeological data are appearing in new types of publications and are finding new avenues for dissemination. The definitions of author, publisher and content creator have become blurred, and entities other than the original excavation team are playing important roles. A close collaboration is required between the excavators and the digital designer and publisher to produce text, graphics and organizational layouts. New formats and presentations are so different from traditional print-based publishing that new techniques must emerge for crediting authors and illustrators, peer review and bibliographic citations. This paper addresses some of the changes that digital media bring to the process of archeological publishing.

Imagine walking through a virtual reality re-creation of an ancient site (Figure 1) — a true three-dimensional space, full of sounds, activities, people, furniture, artifacts and architecture that you can experience as if you were really there. When you see an artifact or a piece of wall decoration that interests you, or you have questions about an assemblage, you can either click on the object (Figure 2) to retrieve instantly a collection of information such as drawings, explanatory text or high-resolution models, or you can activate a search window and query a database that contains all the three-dimensional models, photographs, excavation notebook pages and text about the site. The results (Figure 3) will be automatically formatted for you into a temporary document that you can read, print, save to disk or take with you as you continue to stroll through the virtual world. This is a glimpse of the future excavation report; indeed such reports are being created now.

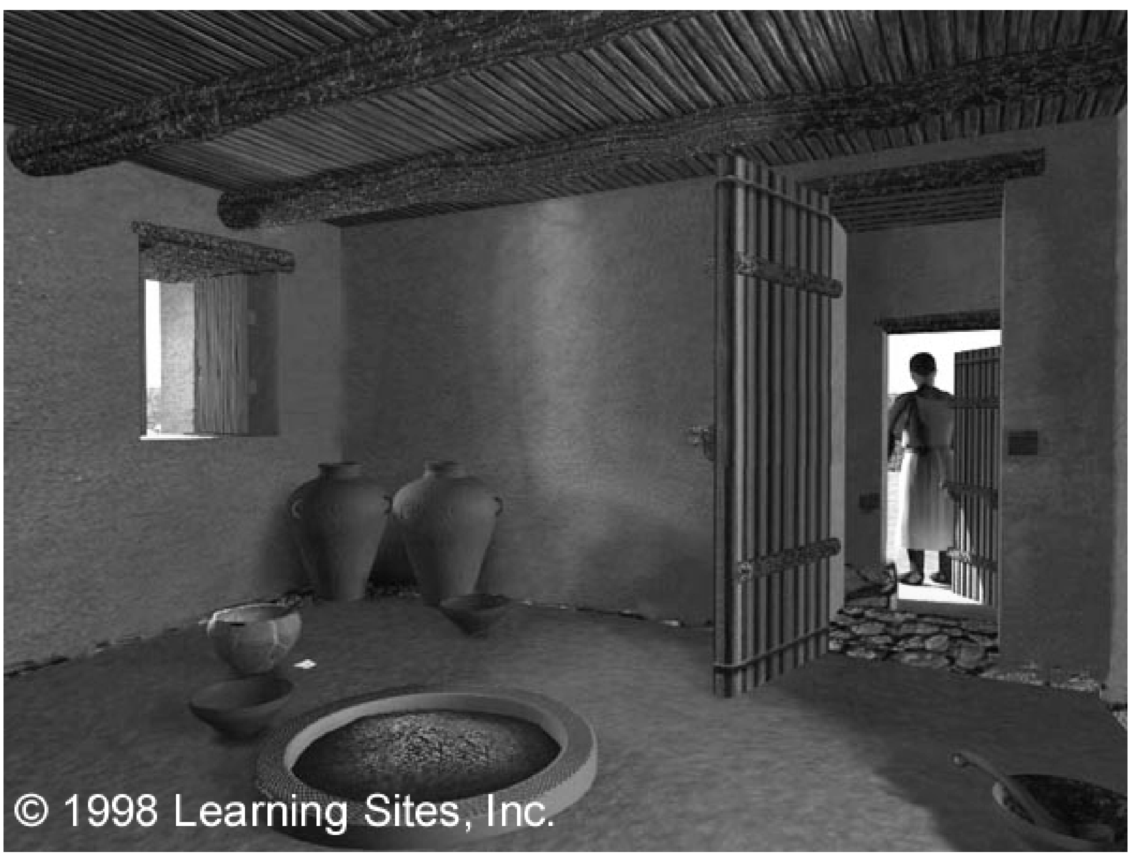


Figure 1. House A interior; rendering from the Learning Sites computer reconstruction of the Bronze Age corridor house from Tsoungiza, ancient Nemea, Greece.

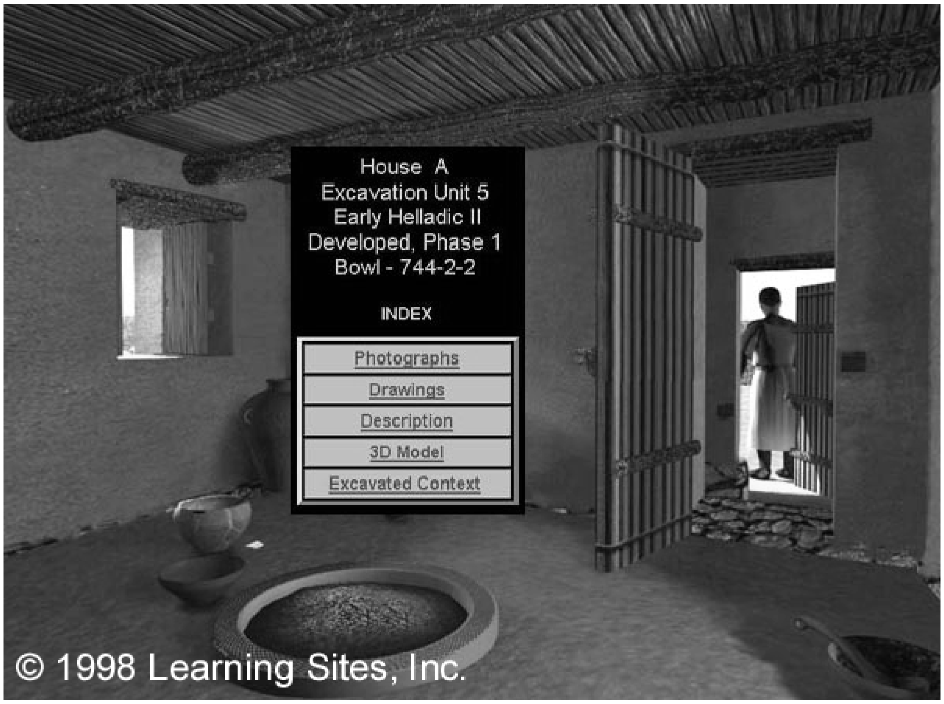


Figure 2. House A interior with local index to bowl; composite image showing elements from the Learning Sites all-digital virtual-reality-based site report of the Bryn Mawr excavations of the Bronze Age settlement on Tsoungiza, ancient Nemea, Greece.

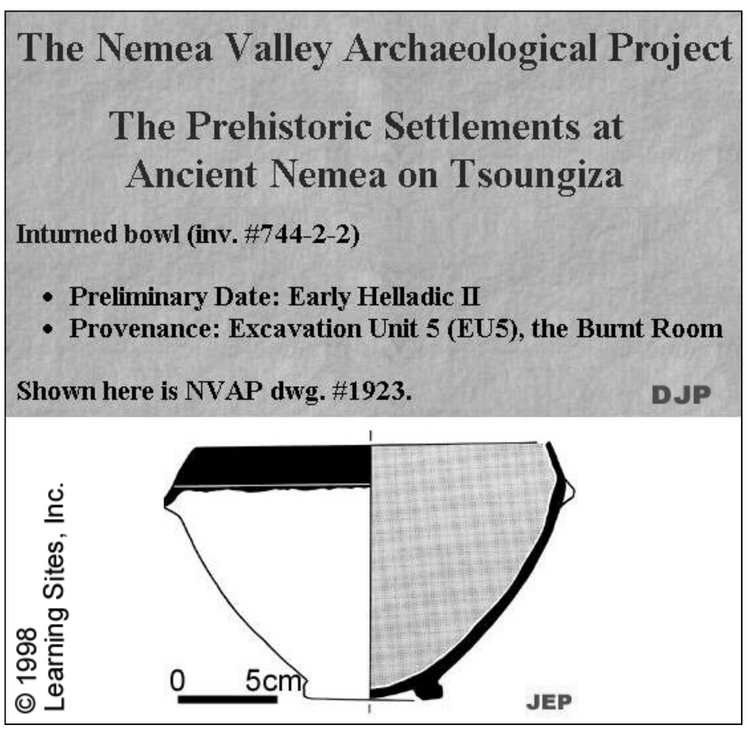


Figure 3. House A, bowl 744-2-2; composite image showing some of the elements accessible directly from the local index of any object appearing in the all-digital virtual-reality-based excavation report of the Bronze Age settlements on Tsoungiza, ancient Nemea, Greece.

Traditionally, information came in discrete neatly defined blocks — books, chapters, and pages — each with an assigned author. In the example above, however, information comes as linked segments of information dynamically created as a result of specific queries. The three-dimensional space — the virtual world — beyond being a window to the past, is a visual index to all the information published via the re-creation, both accessible and changeable depending on how it is used.

This scenario of information retrieval raises some questions. Who is the author of the document retrieved as a result of your search? Who is the author of the particular scene in the virtual world you are experiencing? Who is responsible for content in such a nonlinear, hyperlinked, multimedia publication, with interactive documents created on the fly as a result of individualized queries? Questions about authorship inevitably lead us to question the definition of a "document," a "chapter" or a "page." Are such divisions meaningful in electronic publications that contain animations, sound and virtual reality? Is the term "author" meaningful here?

What happens when a third party — not part of the excavation team and not a publisher in the traditional sense — takes on the roles of both content creator and publisher working in close collaboration with archeologists? How are the traditional definitions of and assumptions about authorship and scholarship affected by this new relationship? What happens to the bibliographic citation conventions?

This paper reviews definitions of "author", discusses how limitations of traditional paper-based archeological publications have encouraged the rise of digital publishing, discusses how Learning Sites — one of the third-party companies actively producing alternative archeological publications — handles its role of content creator and concludes with some issues raised by the movement towards an all-electronic information universe. Offered here are some possible solutions for assigning and citing authorship in an increasingly digital world.

Definitions of Author

The definitions of author, publisher and content creator blur when entities other than the original excavation team play important roles in the process of data analysis, organization and dissemination. New nonlinear, hyperlinked, multimedia publications bear so little resemblance to traditional monographs that it may be necessary to develop new techniques for such things as crediting authors and illustrators, peer reviewing the results, and citing all or parts of the final work.

The topic of authorship for electronic media only now is coming to the attention of standards organizations and professional societies. For example, "author" is not one of the basic elements in the Dublin Core — a set of categories or metadata for the description of all kinds of textual and image resources, especially electronic ones. Instead, the Dublin Core has chosen the term "creator" and defines creator as "the person(s) or organization(s) primarily responsible for creating the intellectual content of the resource."[1]

On the other hand, the Institute for Scientific Information's Hypertext Terminology and Concept Glossary[2] defines "author" as the "writer of an article, chapter or other complete work." There may be multiple authors, and they may be individuals or organizations, but for cataloguing a work, one must be chosen as the primary or senior author.

The University of Texas Southwestern[3] has proposed graphic standards and electronic publication policies in which "document" is defined as the basic element for digital publications. In this context, documents mean text, audio, video, graphics and similar information. Although not precisely defined, the standards require that authorship must be indicated on the document.

Publication Types

Traditional archeological publications are prepared in six steps: data collection and analysis, writing, publisher or journal review, editing, publishing and critical review. This pattern has been fairly stable in archeology for two hundred or so years. Each step has specific individuals responsible for specific portions of the whole — specific paragraphs, chapters, appendices and illustrations — and for aspects of the final product such as editing, printing and publishing.

Recently, discontent with this process has contributed to the popularity of electronic publications. The digital environment, perhaps, can mitigate, if not eliminate, perceived problems with traditional methods, such as the high costs of production, the small sample of excavated data that can be published, the difficulty in updating a work on paper when new data or syntheses appear and the high production and distribution costs that limit the audience who can have access to the work.

Electronic publications promise significant improvements by delivering works to a wider audience faster, more efficiently, and with vastly more data and analysis than possible with traditional paper-based publishing. Digital publications are easy to update; they are inexpensive per unit of information to produce and distribute; the end results can be interactive, multi-user, and customizable; and they can offer a vast amount of up-to-the-minute information to the public.

Medicine, high-energy physics and history of technology publishing, for the most part, eschews monographs and instead, prefers the dissemination of scholarly work via articles — and largely in electronic journals. Because most electronic publications in the sciences and elsewhere, however, have tended to be organized like normal text-based works but published in digital form, there is little precedent for the developments outlined here with regard to archeology. The changes described here for our multimedia, virtual-reality-based reporting have serious implications for the way we perceive, cite, and use archeological data.

How is Learning Sites, Inc. Involved?

Just what these changes are and how they will affect our research and publication will be come clearer after a review of a few examples of how Learning Sites' staff — with expertise in archeology, architecture, architectural history and information science — collaborates with clients.

For the Assyrian palace of Ashur-nasir-pal II, at Nimrud (Figure 4), we fully participate with an international team of specialists to discuss each reinterpretation of the architecture, sculptural program, decoration, history, and use of the monument. The final monograph, an exceptional research resource, will integrate intelligent agents programmed to lead researchers through the massive amounts of data being collected and presented entirely in virtual reality.

For the religious center of Gebel Barkal, ancient Nubia (Figure 5), we work closely with curators at the Museum of Fine Arts, Boston, examining primary documents, debating various reconstructions of the architecture and the wall decoration and collaborating in the preparation of analytical text, as we move toward a site-wide electronic report of the current excavations.

Finally, for the Bronze Age settlement of Tsoungiza, Ancient Nemea, Greece (Figure 6), our staff works closely with the excavation leaders to study the architectural remains and collaborate in discussions about construction techniques, use of materials and design for each reconstruction (Figure 7). We also are reworking the excavators' database into a searchable front-end for our virtual reality re-creations of each trench (Figure 8) resulting in an entirely new mode of studying the site, which will integrate all the excavated information including photographs, drawings, field notebook pages and analyses.

In each case, the three-dimensional modeling techniques and the virtual reality environment led us to new insights about artifacts and buildings and their use that could never have been realized using traditional static, two-dimensional visualization or presentation methods. For each project, we are careful to record who was responsible for each decision, each item of text, each image and each aspect of the three-dimensional models.

When companies like Learning Sites or the Digital Archaeology Laboratory at UCLA step in to organize primary data into one of these exciting new publications, they transform the data into interactive bits that defy traditional chapter designations, or change the look and even the content as it is recast into the hypertext, multimedia final work. The presentation format, the organizational format, the presentation media and the visualizations that support, supplement and often contain excavation material are largely our creation. These new third parties, while not adding raw data, are responsible for a good portion of the content and are responsible for taking data and analyses and presenting them in ways that offer new insights not possible in linear, static, codex-based works.

Consequences of New E-formats

How then, can we assign authorship to each interactive bit and to any given constellation of interactive bits? Before concluding with some suggestions, let me enumerate some other consequences of these new electronic works.

How does one peer-review a hypertexted, interactive, nonlinear publication, which will appear different to each reader, depending on the paths chosen or searches performed? What is being reviewed — the text, the links, the interface, the visualizations, the entire package? When virtual reality becomes the container and medium of navigation between the written word, the static image, the moving image and the interactive three-dimensional environment, will the reviewer or editor be fluent enough in all four technologies, in all four interfaces, to move fluidly among computer-based datasets? Further, and more fundamentally, how will reviewers, editors or other scholars know who authored what piece of the whole when there are no chapters, no pages, no neat packages of linear text-only information. And, consequently, how does one cite such a work or a piece of such a work? How does one ascribe an author to a multimedia three-dimensional dynamic screenful of information? Does not critical inquiry rest on being able to present replicable substantiation for arguments? Don't scholars need to know who authored what opinion so that we can relate this information to the wider body of knowledge by the same person in order to judge the writer's credibility? If scholarly argument is based on attribution, then don't we need to be able to cite a person and a specific location from which we got our material?

There are no scholarly precedents for what is happening here. Yes, there are lots of multimedia and even multi-author CDs out there, but since they do not purport to be scholarly publications, there is no pressing need to cite specific information or a single "page" or identify a specific author's contribution to the final work.

The Association for Computing Machinery has addressed the issue of evaluating e-publications when reviewing for tenure, an important point for their audience because ACM soon will publish only in

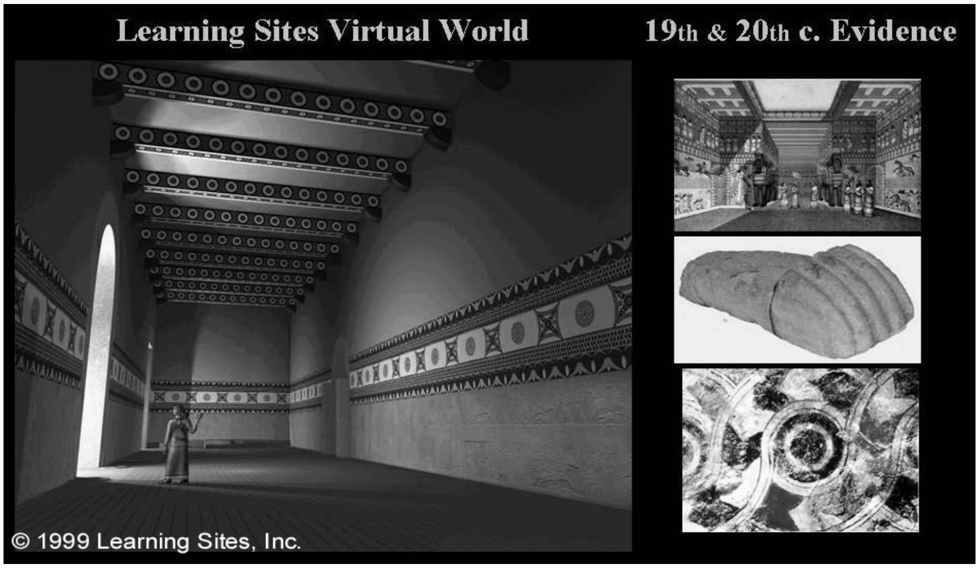


Figure 4. Northwest Palace of Ashur-nasir-pal II, Nimrud, ancient Assyria. Left image: rendering of the Throne Room. Right images – evidence used as the basis for the reconstruction: top – A. H. Layard's 19th-century reconstruction of a similar space; center – hand corbel found in the excavation of the Throne Room; bottom – fragment of plaster unearthed in the Throne Room.

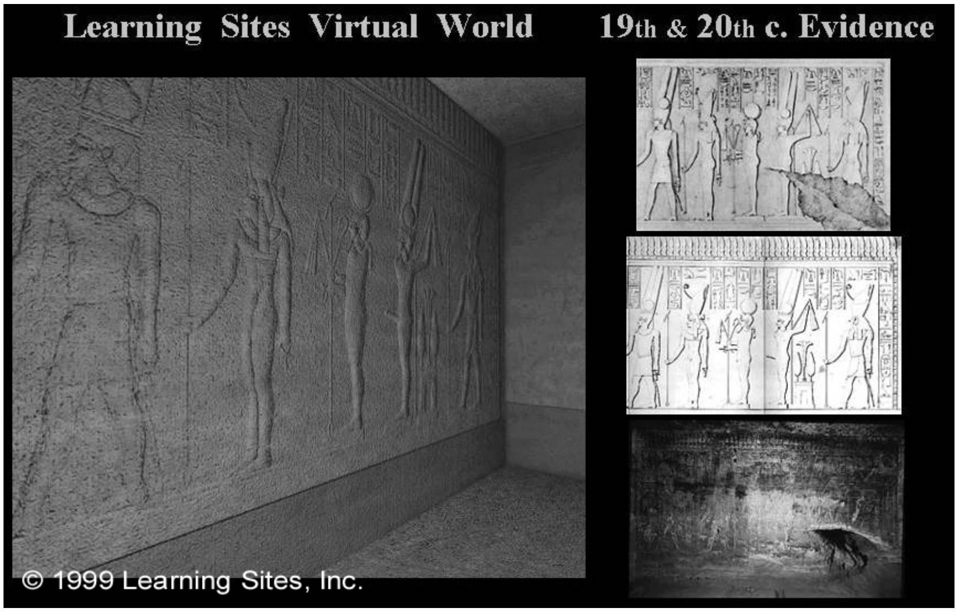


Figure 5. The Temple of Mut (B300), Gebel Barkal, ancient Nubia. Left image: rendering of a wall in the sanctuary. Right images – evidence used as the basis for the reconstruction: top – drawing of the wall decoration by M. Linant de Bellefonds, 1822; center – drawing of the same wall decoration by Karl Richard Lepsius 1844; bottom – photograph of the remains of the wall decoration by Enrico Ferorelli, 1984.

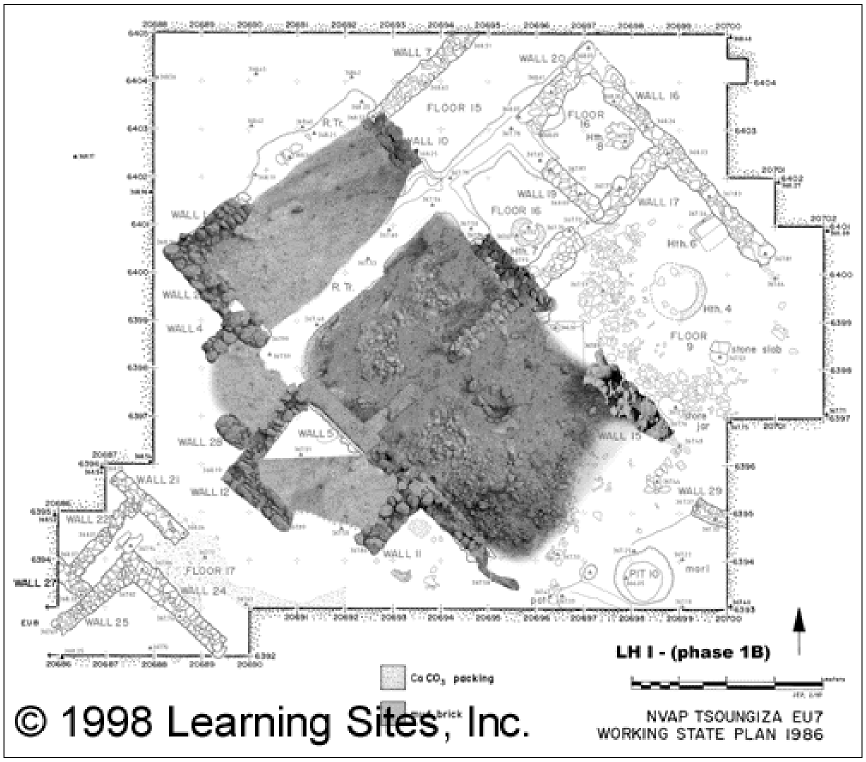


Figure 6. Late Helladic I (LH I) phase plan from Excavation Unit 7 (EU7), Tsoungiza, ancient Nemea, Greece. Phase plan showing the state plan drawing of EU7 overlaid with photographs taken of the actual remains. Each phase plan image is linked using Javascript programming to the photograph database.

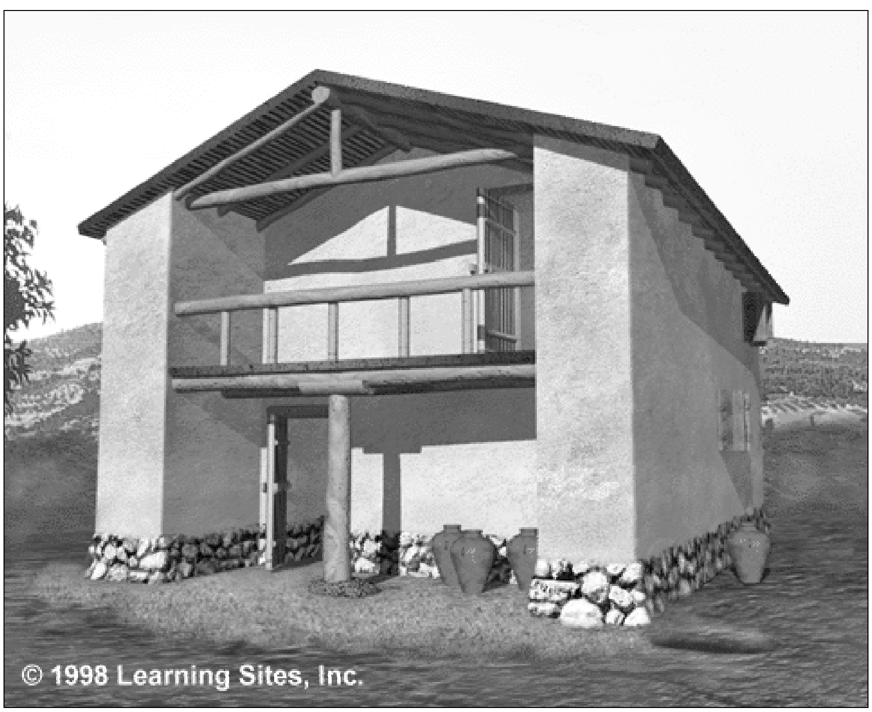


Figure 7. House A exterior; each aspect of the rendering has a Javascript routine linked to it so that when the cursor is dragged over an element, a flyout identifies the feature, explains whether it survives or is conjectural, and links the user to the photograph database if the feature was found during excavation.

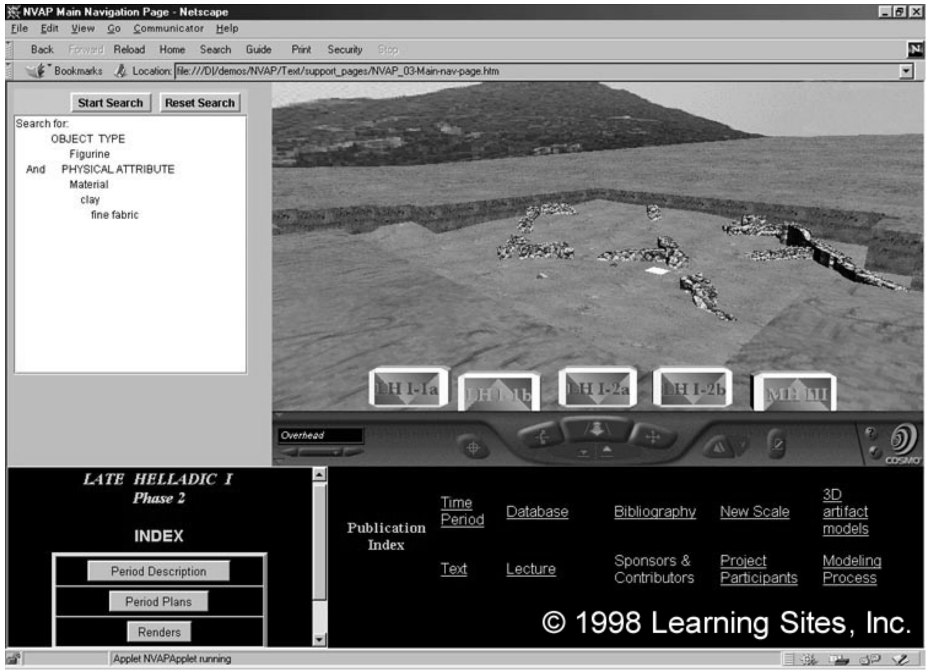


Figure 8. Sample screen grab from the preliminary user interface. Top left – Java-based search engine linked to the entire database of finds; top right – the image window, currently showing the virtual reality reconstruction of Excavation Unit 7; lower left – the Local Index frame, currently set to items available for the Late Helladic I, Phase 2, period of the site; lower right – the General Index to the publication.

electronic format. ACM assures readers that “traditional criteria and standards for appointing editorial boards and refereeing papers [and] warrants that scientific papers published electronically in ACM refereed journals meet traditional scientific and engineering standards and should be accorded equal stature with print publications”.[4] ACM, however, is referring mainly to papers that are organized like traditional printed text, though stored and accessed digitally.

The interactive, interwoven resources being created by Learning Sites and others for the benefit of archeology are unique. The questions raised here are new; they have been addressed neither by such organizations as the Modern Language Association, the International Intellectual Property Alliance, the European Community’s Information Society Project reports, the International Organization for Standardization, nor in the *Journal of Electronic Publishing* or in the Arts and Humanities Data Service standards.

Conclusion

Electronic publications produced, for example, by Learning Sites are not jigsaw puzzles in which there is a single unity of all the myriad bits of information at any stage of its use. One cannot cite an author of a single page because the page or screen of information is dynamic and variable.

One key to the attribution of authorship may be to accept e-publications as simply not the same as codex-based works. Digital publications have many different formats, encompass a different amount of data, provide different methods of presentation and are not linear. Maybe we cannot merely take existing citation and author definitions, created and refined for the codex era, and assume they will work for electronic media. Maybe we need a different paradigm.

To develop a new paradigm, we could look for analogies elsewhere. For instance, making a movie — a multimedia affair with sounds, action, lighting, and words — means taking a written work, and adding creative input and changes from a director, producer, lighting and sound engineers, support staff and actors. An entire crew is responsible for the final work; each individual can build on another’s material. No one part can stand alone and the long list of credits for the final work cites each member of the team who may have had very little to do with the original data or content. However, despite the similarities, we are here dealing with different intentions. Movies and other analogous collaborative works are not meant to be scholarly publications and thus need not answer to the same rigor.

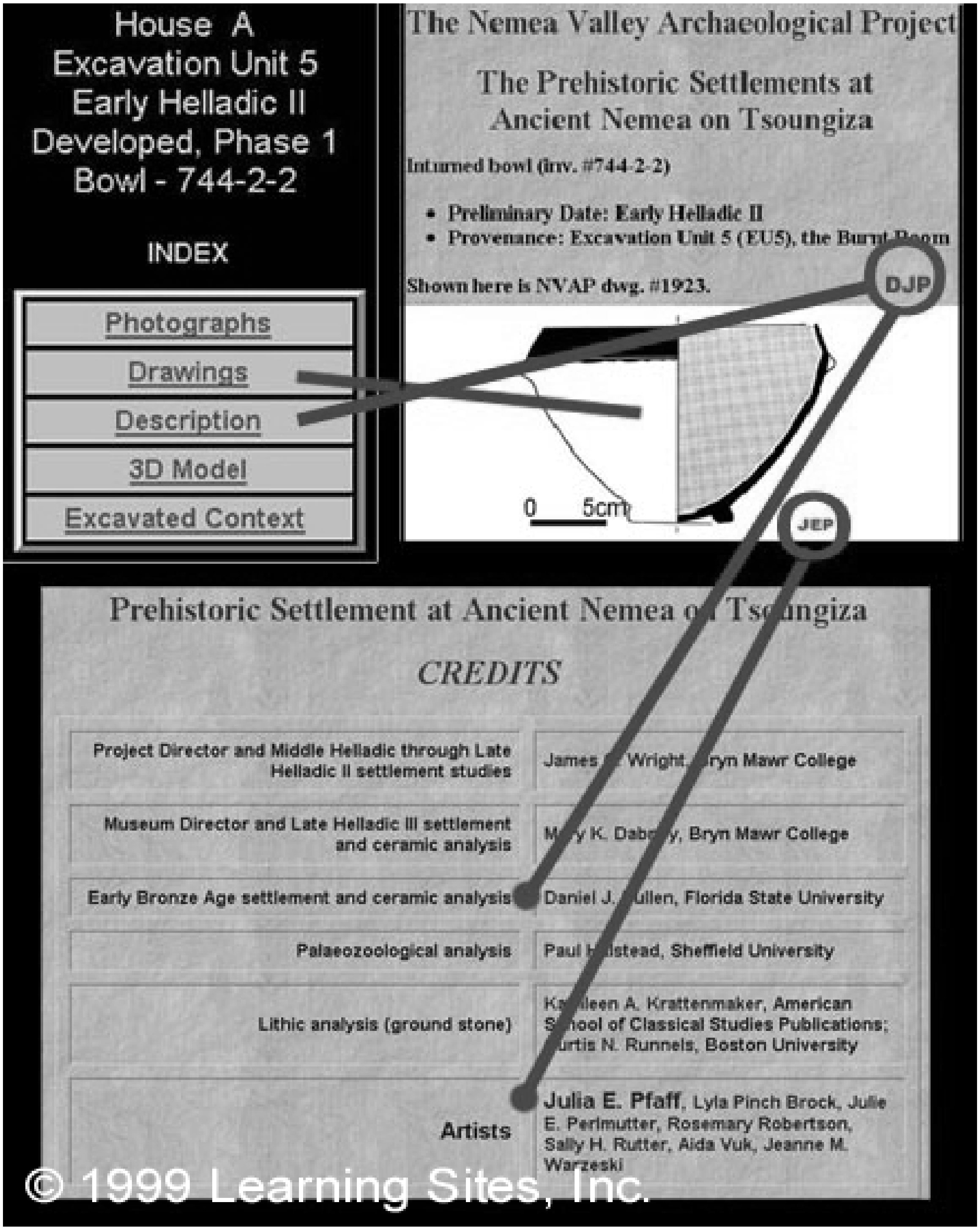


Figure 9. Composite page showing how creator credit information is displayed and linked in the all-digital virtual-reality-based site report of the Bryn Mawr excavations of the Bronze Age settlement on Tsoungiza, ancient Nemea, Greece.

Instead, what if we return to this essay's initial image of the future; that is, jump over the present uncertainties for a moment. Imagine perusing an electronic publication online. You attach a digital bookmark to an interesting location in the virtual world or tag a bit of information that you may want to cite later. The bookmark will automatically contain your name, e-mail address, date of visit and your personal annotation. Remember we are live and online. When other scholars visit the same site and wish to cite the data, they could choose to see all the bookmarks set by other visiting scholars or only their own — as is possible with electronic reviews at Internet bookstores. Researchers will be able to link directly to that bookmark — the citation itself now — which could also contain information about the path, the multimedia environment and the author of each bit of data in view.

Thus, in an entirely electronic environment it is easier to envision how authorship and citations can be handled. Meantime, the difficulty is creating citations to dynamic digital publications from compartmentalized static paper-based ones. We could maintain some of the format of codices in the electronic works. That is, include traditional linear descriptions and analyses in addition to the hypertext multimedia formats. Scholars could use the interactive aspects for research, but reviewers and tenure committees could opt to read the linear text. The problem here is that the linear blocks of text with single authors bear little relation to the information and interconnections presented in the actual digital publications.

One solution that we are implementing in our electronic publications is to tag each snippet of information with the creator's initials (Figure 9). Each piece of text, each image, each virtual reality texture is linked to its creator's background information and level of contribution. Although citations may be cumbersome, at least authorship can be maintained, and critical analyses and tenure reviews can continue with attributions intact. This is all feasible now, and provides a relatively painless transition to a more electronic future when still newer methods of disseminating archeological information appear.

The transition period we are now in will not likely supply the entire solution; things will evolve as we come to grips with emerging technologies and begin fully to take advantage of new options. Nevertheless, there is no reason to stop trying to advance our understanding of cultural history, nor our visualization and presentation of our interpretations. This is an exciting time for our profession, with still untapped resources on the horizon, such as three-dimensional semantic networks, projection holograms and multi-user stereo virtual worlds.

Certainly, in archeology's digital future, there will still be individuals who will write descriptions and analyses, who will be the signatories and whose e-mail will be linked to the result, but the responsibility for the whole, for the content and presentation of the work, will likely reside with a much larger group than is common today. Our study of the past need not rely on methods of the past.

Notes

1. Stuart Weibel. *Discovering Online Resources. The Dublin Core: A Simple Content Description Model for Electronic Resources*. Arts and Humanities Data Service. Nov. 17, 1997. <www.ahds.ac.uk/public/metadata/disc_03.html>.

2. *ISI Hypertext Terminology and Concept Glossary*. Institute of Scientific Information. 1999. <www.isinet.com/help/glossary.html>.

3. *Publishing on the WWW Server at UT Southwestern*. University of Texas Southwestern. Sept. 1, 1997. <www.swmed.edu/internet/publishing/>.

4. *ACM Policy on The Quality of Refereed Electronic Publications*. Association for Computing Machinery. Jan. 15, 1997. <www.acm.org/pubs/quality.html>

Online not In Line:
Geospatial Data, Decision Support and the Internet

JAMES A. FARLEY

Recent innovations in information technology have focused on basic problems of openness and relied on published standards as a basis for developing products and solutions that work together effectively and transparently in the distributed landscape dominated by the Internet. Of the many types of complex data that are being leveraged in this process perhaps none are more significant to archeologists than the geospatial data that fuels Geographic Information Systems applications. This paper examines the range of technologies and strategies that combine to support delivery of a variety of geospatial data and information to the desktops of decision-makers, researchers and policy-makers.

Recent initiatives across the spectrum of information technology have focused on basic problems of openness, access and published standards as a basis for implementing interoperable strategies and products such as Object Management Group[1], W3C[2]. The growth of Web driven technologies such as Java[3], Enterprise Java Beans, eXtensible Markup Language (XML)[4] and related protocols (XSL, etc.) and the infrastructure technologies such as CORBA or DCOM on which they depend reflect this trend. Parallel developments have been underway in the arena of geoprocessing and spatial data management. Under the auspices of the OpenGIS[5] Consortium the geoprocessing industry has come together to collaborate on developing a specification and associated implementations that will alleviate many of the constraints typically associated with geospatial applications.[6]

The proliferation and maturation of these foundation technologies is not a random event. The direction and the pace of growth in these technologies is a tangible response to the needs of users who require maximum access to rich, diverse information and demand that minimal constraints be associated with the processes of locating and consuming data and producing new information.

These are the forces that are driving technology markets today: openness, access, interoperability and published standards; delivering the right data to the right consumer in a timely manner with a minimum of overhead and intervention; creating an environment that supports the delivery of raw data; derived products and highly customized information commodities relying on an extensible interface that supports a diverse community of users and serving education, government and business at all levels from a common information base.

In this environment, outside of the domain of e-commerce that continues to grow at an exponential rate, perhaps no user communities or domains have a greater need for technology exhibiting these attributes and these capabilities than government and academia. In these communities a wide range of archeological initiatives are sponsored, fostered, undertaken and administered. The inherently distributed nature of many projects and the diverse composition of the community indicate that a broad cross-section of the archeological community itself stands to derive significant benefit from many of the changes that are under way or imminent.

Government and academia at all levels have consistently cultivated and embraced technology in a search for solutions that complement their hierarchical or distributed structure, their heterogeneous business model and their mission. This is particularly true in the case of spatial technologies that have the capacity to play such an integral role in much of the work of government and have come to be widely acknowledged as instrumental tools in innovative archeological applications.[7]

Definitions

There are a number of basic concepts that form the foundation for this paper. Those that are most central to this discussion are described below.

OpenGIS — a consensus driven process that is directed towards bringing broad-based representatives from industry together with diverse user communities to:

- identify and describe geospatial application requirements,
- capture requirements as workflows and use-case scenarios,
- formalize workflows and use-case scenarios as interoperable interface specification and
- facilitate the validation of specifications through the promotion of implementations, testbeds and interoperability initiatives.

Spatially Enabled Warehouse — an aggregation of broad-based general purpose and specific data that includes geometry defining the location of features, events and phenomena on the surface of the earth. The warehouse also facilitates user inquiries and data mining via a compendium of business rules and domain logic that promote information extraction. This library of business logic that assists the user in locating and filtering information will grow in size and become increasingly sophisticated over time.

OpenGIS Technologies — Commercial-Off-The-Shelf products that are built in accordance with the specification developed in the Open GIS Consortium's Technical Committee.[8] Certification and conformance with the specification is managed under a formalized set of policies and procedures administered through the consortium.

Enterprise Community — A definable, distributed community such as academia and government as an integrated, corporate entity or a mechanism that is composed of distributed, interdependent components that share corporate information in a seamless technology environment to improve decision making.

Seamless Warehouse of Arkansas Geodata

The University of Arkansas, Center for Advanced Spatial Technologies has been in the business of delivering digital geospatial data, geoprocessing resources and products derived from digital geodata to users in government for nearly ten years. In fact, technology transfer in the area of geospatial applications is a central component in CAST's mission. CAST is a founding member of the OpenGIS Consortium and has a keen interest in the progressive use of distributed technologies as an infrastructure in support of interoperable GIS and the delivery of geospatial content to desktops in state and local government and across the academic community. Over time we have become sensitive to the role of geodata in these application contexts and to the potential this data might have to improve decision-making at multiple levels and in many settings. We also have become convinced that a major component or class of digital geodata known as "Framework data"[9] — the baseline products commonly associated with many geospatial applications such as elevation, hydrology, transportation, political boundaries, soils and vegetation — is actually a corporate resource in the enterprise computing sense of the term. In other words they are likely to be required on a regular basis by a large percentage of the "enterprise" as part of the day-to-day business. As such these data are most effectively managed as a shared resource available across the entire enterprise in a form that has integrity and is readily consumable.

Building on technology trends and in recognition of the enterprise community's needs, CAST, collaborators from across government and a team of private sector partners is assembling a comprehensive, spatially-enabled data warehouse that will deliver rich, complex data to users in state and local government, K-12 programs and a range of other clients throughout the state. The Seamless Warehouse of Arkansas Geodata is being constructed using spatially enabled, object-relational database technology and industry sanctioned OpenGIS interfaces that support access from multiple, heterogeneous client-side applications. SWAG is built on an innovative model that relies on the Internet, inherently distributed relationships and technologies, open standards and a series of private-public partnerships to address existing problems and deliver new services.

When completed, SWAG will exceed one terabyte in size and will house metadata, attributes and spatial geometry for vector and raster data and will support data delivery, data mining and data warehousing applications via domain specific, spatial middle-ware. SWAG will provide a uniform view of geodata products that can be located, evaluated and delivered over existing network infrastructure to end-users and application specialists. The data will be housed in an Oracle object-relational warehouse capable of managing a wide variety of complex data in a manner endorsed by both the industry and international standards organizations. The data will be available to users over the Web via OpenGIS-compliant interfaces creating the vehicle for interoperability. Because of this, end-users — regardless of the GIS software they use — will all have access to the same baseline data to support their applications which will significantly reduce the "buy-in" cost associated with GIS-startup at offices in state and local government. The more profound impact realized over time will be an environment in which information drawn from a common database results in better decisions.

Objectives

SWAG has a series of well-defined objectives:

- Construction of generic, prototype mechanisms for collecting, managing, publishing and distributing a range of geospatial data and information products, and the metadata used to describe these products.
- Reliance on both the statewide telecommunications network using standard Internet technology and the emerging Internet II architecture with its broad bandwidth backbone as the infrastructure that supports these activities.
- Support for access to the seamless warehouse repository using unmodified browsers and

heterogeneous Commercial-Off-The-Shelf client-side applications from a range of industry leading vendors such as Autodesk, Environmental Systems Research Institute, Intergraph Inc., MapInfo Inc., etc.

- Conformance with OpenGIS, International Standards Organization and Federal Geographic Data Committee standards relating to metadata and geodata

The Technology Mix

The viability of any solution for packaging and distributing information depends, at least partially, on the technological landscape for which it is designed and crafted. The technology framework that will provide the engine, the delivery mechanism and support the warehouse and data mining functionality for SWAG is consistent with current trends that are evident across the information technology community. By blending the best components of conventional IT philosophy with emerging strategic technologies, SWAG should be stable, immediately competitive and useful into the future. Given the rate, pace and directions of technology growth, change and extinction, these are good attributes to be able to ascribe to any system that is scheduled for implementation and deployment.

There are five major components of the SWAG architecture, each of which is essential to the successful development and deployment of the system (Figure 1).

1. Object-Relational Database (ORDB) technology
2. Unix Enterprise-class computer
3. High bandwidth network connectivity
4. A range of highly structured, standards compliant technology
5. Multiple, user-selected client-side applications

This architecture is consistent with many engines that are designed to respond to user requests and push information onto the stateless platform provided by the Web. In this case a Unix enterprise computer provides the platform for a robust data management engine which sits on a high bandwidth network exposing metadata and data through published, standards compliant interfaces to software which runs locally on a remote user's desktop machine.

The enterprise Unix server for SWAG is an Enterprise 5000 Data Warehouse from Sun Microsystems. Data management is handled by Oracle Corporations Oracle Enterprise Server 8i database engine using an object-relational data model for optimal performance in an environment characterized by geometric objects managed using the spatial cartridge technology. Standards-compliant interface technology is provided via Oracle's implementation of the OpenGIS specification for open geoprocessing. The network bandwidth is provided via Bay Networks ATM technology (OC3/12) that provides connectivity with both the statewide backbone and the National Science Foundation's Internet II.

In each case, the foundation technology — server, database engine, application development environment and the ATM connectivity and switching capacity — has been provided at little or no cost by a cooperating vendor for the express purpose of participating in the collaborative partnership envisioned for building and fielding SWAG in service of the broad target enterprise community.

The SWAG initiative will result in a common warehouse of geospatial data products that is accessible from a range of common desktop platforms that have access to the Internet. When completed, SWAG will be a viable resource available to a wide range of government agencies and educational entities in the state. By consolidating a range of complex map and attribute data in a common warehouse that is universally accessible, SWAG will enable a range of applications that have been unavailable to many throughout the state prior to this time. A generalized model of the SWAG initiative is presented below (Figure 2).

SWAG will continue to evolve as a collaborative partnership that spans academia, the private sector and many elements of government. The basic framework for this collaborative already has been established with a firm network of partners committed to promoting and refining the basic technologies embraced by SWAG. Given this, SWAG is a resource that will provide support and benefit to government and education in Arkansas for years to come. The project design will enable us to establish the baseline technologies and bring them online so that SWAG can grow and expand as more partners enter the collaborative.

At this juncture it is important only that SWAG has a specific focus and a target for its activities and growth. The focus for SWAG is technology transfer with an agenda that enables new solutions by leveraging proven, mature technologies. The target for SWAG is "real-world" applications in state and local government such as planning, disaster recovery, emergency response, monitoring health statistics — birth, mortality and morbidity — mapping in support of legislation and education and a host of other spatially dependent applications.

SWAG and the Enterprise Community

SWAG is creating a number of opportunities and providing users with a range of functionality. In particular SWAG will:

1. Create a seamless repository of spatial data of both vector and raster type that is maintained at multiple scales ranging from local very small scale — such as 1:1,000 — to 1:12000 to 1:500,000 or higher and at a number of resolutions from 5 meter to

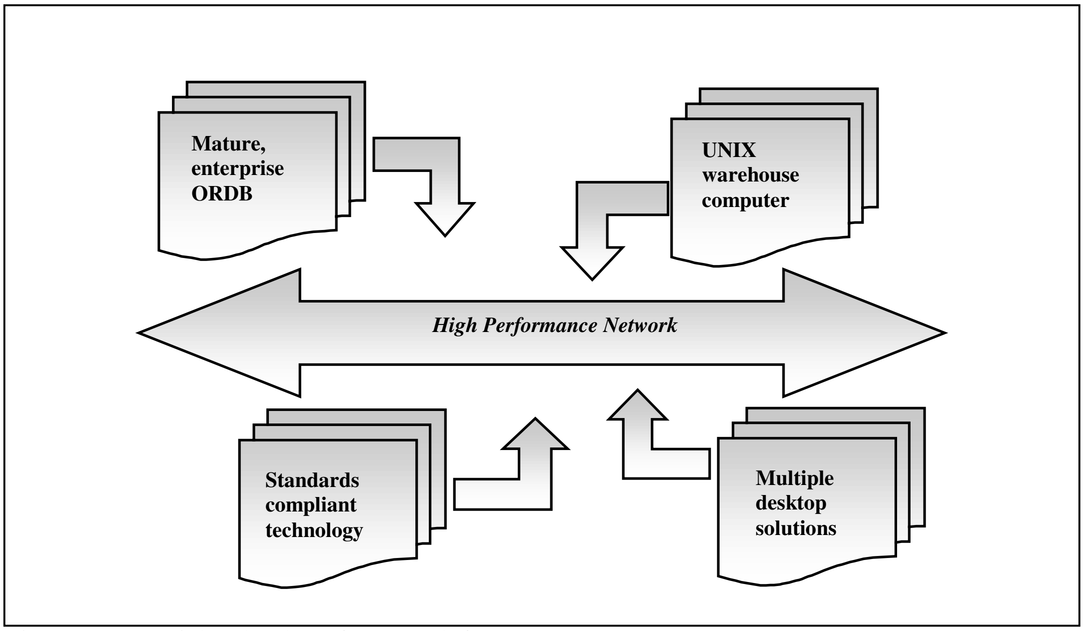


Figure 1. Essential technology framework for SWAG.

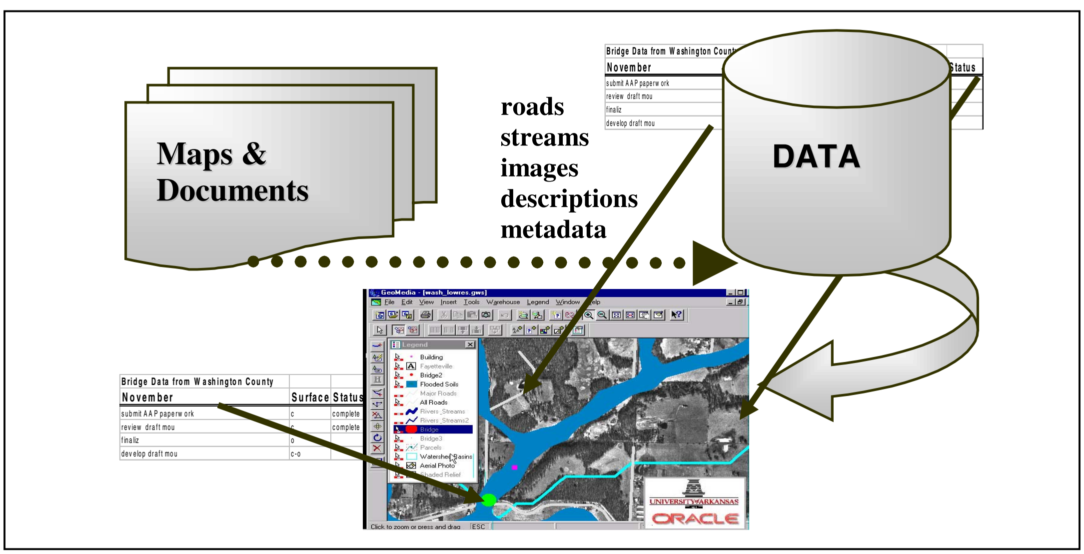


Figure 2. Schematic representation of SWAG and its application.

greater than 30 meter. In this seamless repository, users will be able to identify and acquire the specific data that they wish to have without the normal encumbrances — file based processing, import, export, clip, etc. — that are associated with file based data transfer.

2. Create a stable infrastructure, firmly grounded in international and industry standards, for storing, publishing and distributing the important digital geospatial resources for the state of Arkansas. By maintaining a proactive involvement in the OpenGIS community and other peripheral technology forums — ISO/SQL3[10], OMG — we will ensure that SWAG and our state's important digital geospatial resources are maintained in a secure, extensible archive that is both open and compliant with general industry trends and directions.
3. Provide nearly universal access to uniform data sets via a published Application Programming Interface (API) based on the OpenGIS Consortium's specification for interoperable geoprocessing which has been adopted by leading industry vendors. Given the wholesale endorsement of the emerging OGIS specification, SWAG can capture a single instance of a geodata theme — TIGER roads, political boundaries, USDA-NRCS soils, hydrologic units or most any other vector or raster product — and expose that one version of the data to any compliant application or browser via a certified interface.
4. Eliminate data redundancy and minimize data translation. These features in turn create the possibility of eliminating data redundancy and costly steps associated with data translation and data import by delivering complex data directly into applications.
5. Move analysts and decision-makers closer to the data resulting in better decisions. By eliminating many data processing steps associated with file based import and format conversion the possibility of extending expert systems and other rule-based decision support applications is enhanced.
6. Protect expensive geodata using basic checks and balances. A robust data management backend supports basic transaction checks and balances to manage integrity and versions. A form of "Decision Insurance" for application specialists who base their decisions on data acquired from the SWAG warehouse is created.

The basic architecture that addresses these needs and delivers this functionality is portrayed in Figure 3. A number of different agencies are shown accessing spatial data from a common repository. It is possible, even probable, that each agency may have different requirements that have resulted in distinctly different hardware and software profiles at individual locations. Historically, this difference in application software would preclude the agencies from relying on a single source of digital data without obtaining costly, often unreliable conversion software. This process of conversion leads to errors and a number of redundant, often inconsistent data sets being created at different locations. The overhead in terms of people time and a general lack of confidence in the baseline data can be significant. For instance, it is not uncommon for each entity in a region that uses geospatial data to have its own "local version" of many of the framework data themes — roads, streams, elevation, political boundaries, vegetation, etc. This duplication is expensive and may lead to inaccurate comparisons based on data sets that are out-of-phase as a result of local modification or incompatible version management.

However, with SWAG different local software applications are reconciled via a common language — the OpenGIS interface specification, Application Programming Interface — that is used to access the data in a single, corporate shared data store. As a result, each of the entities accessing the database can benefit and draw information from SWAG. As the SWAG product and the OpenGIS specification mature, the dashed lines that connect an agency or an application to the warehouse will actually become rule-based vehicles for mining information based on the domain rules and business requirements associated with individual agencies and entities.

Looking Down the Road

The SWAG system is still in a formative state. Working in conjunction with NASA, the Federal Geographic Data Committee and the Arkansas State Department of Information Services the database is being populated with the basic corporate data themes and the metadata that are needed to field the product across the region. In addition to populating the warehouse with corporate data we are experimenting with a range of server-side software or middleware that will help enable client-side applications and connections via standard browser technology. Middleware will provide the foundation or the point of departure for developing implementations that embody specific business rules associated with end user communities. For instance, if a user, say a forester, has an established routine associated with a specific area, particular data sets, a structured set of questions or even a set of logic (rules) this information may be retained on the server side. Once they are preserved in this manner these preferences or rules may be invoked by the user as a customized service that is made available when they access the system. Once the basic compendium of "framework-type" data has been assembled the identification, definition and implementation of this rule-based middleware will

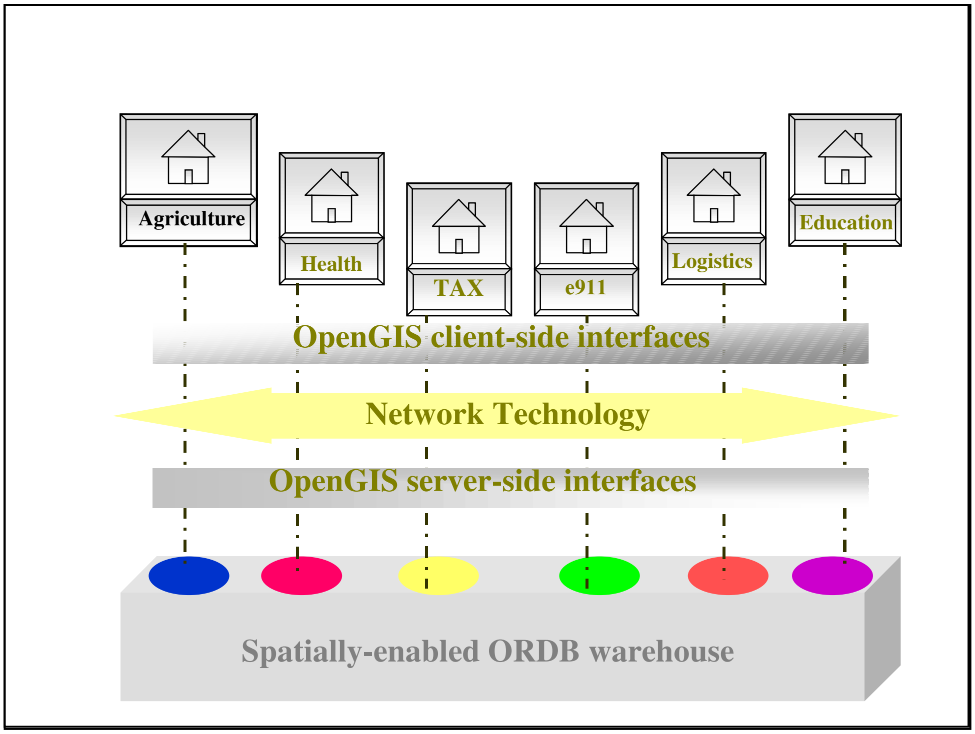


Figure 3. SWAG in support of the enterprise community.

become the long-term, ongoing work of the project.

Summary

As our ability to capture and distribute complex information grows we should re-examine the models used to manage and disseminate this data. The SWAG project is an experiment in creating, storing and using geospatial data that acknowledges significant growth and evolution in database technology, networking capabilities and interoperability. In particular, SWAG is an exercise in integrating emerging technologies based on open-systems standards to construct a framework that helps drive new approaches to information management and decision making in a broad cross-section of the public sector and the academic community. In many respects this endeavor is as much a proof of concept for an institutional or enterprise-type approach to information management and dissemination as it is a yardstick used to measure the growth of open standards for distributed geoprocessing. With SWAG, OpenGIS moves out of the consortium and into the real world to provide the infrastructure that is needed to support the enterprise community.

Acknowledgments

The SWAG project is supported by many collaborators, including Michael Hipp and Susan Cromwell of the Arkansas Department of Information Services; Alex Tuyahov of the National Aeronautics and Space Administration; Doug Nebert and the Federal Geographic Data Committee; Frank McManamon, Veletta Canouts and Terry Childs of the National Park Service-Archeological Assistance Division; Jack Pellici, Michael Fisher and Steve Cooperman of the Oracle Corporation; Michael Cosentino of SUN Microsystems; Larry Ayers, Tom Clemons, Susan Nolen and Preatha Pulasani of Intergraph Corporation; Steve Talbot and Sandra Johnson of MapInfo Incorporated; Dr. Robert Moses of PCI Geomatics and David Maguire, Mike Phoenix and Ed Crane of Environmental Systems Research Institute.

Notes

1. *The Common Object Request Broker Architecture* (Framingham, MA: Object Management Group, 1991).
Object Management Architecture Guide (Framingham, MA: Object Management Group, 1992).
Object Management Group.1999. <www.omg.org>.

2. World Wide Web Consortium (W3C). 1999. <www.w3c.org>.

3. <java.sun.com>.

4. <www.w3c.org>.

5. OpenGIS is a registered trademark of the Open GIS Consortium.

6. Kurt Buehler and Lance McKee, eds., *OpenGIS Guide Introduction to Interoperable Geoprocessing*, 3rd Edition (Wayland, MA: OpenGIS Consortium, 1999).

OpenGIS Consortium. 1999. <www.opengis.org>.

7. K.M.S. Allen, Stanton Green and Ezra B. Zubrow, *Interpreting Space: GIS and Archeology* (London: Taylor and Francis, 1991).

Herbert Maschner, ed., *New Methods, Old Problems: Geographic Information Systems in Modern Archeological Research* (Carbondale: Center for Archeological Investigations, Southern Illinois University at Carbondale, 1996).

8. *OpenGIS Abstract Specification* (Wayland, MA: OpenGIS Consortium Technical Committee, 1999).

OpenGIS Implementation Specification (Wayland, MA: OpenGIS Consortium Technical Committee, 1999).

9. Federal Geographic Data Committee. 1999. <www.fgdc.gov>.

10. International Standards Organization (ISO). 1999. <www.iso.ch>.

Surfing Indoors: Bringing the Net into the Classroom

DAVID L. CARLSON

As the rapid growth of the Internet continues, opportunities to use it to enhance teaching about archeology have expanded as well. While flashy Web sites help to boost students' interest in archeology, there are a variety of simple ways to use the Internet that do not involve complex programming skills. Class Web pages and electronic conferences are good ways to begin. This paper explores a variety of ways that instructors can use Web pages and e-mail to stimulate interest, deliver course materials, facilitate communication and develop critical skills.

While field archeology has changed dramatically in the last ten years with the introduction of global positioning systems and electronic distance metering, so also has our ability to communicate research results to students and the public. Most of us are still struggling to keep up with new developments while we seek ways to use the new technologies to best advantage. This paper focuses on how the Internet and the World Wide Web fit into the teaching of archeology and anthropology.

In the strictest sense, this paper is misnamed. It is not about bringing a computer to class so that your students can gather around it like a campfire while you surf exotic Web sites. Instead, it is about ways that you can use the Web to expand your existing arsenal of teaching tools.

Some of the most valuable contributions the Internet makes to your teaching are pretty mundane on the surface. They are also the easiest to learn and the easiest to incorporate into your courses without dramatic changes in content or presentation. This paper also discusses methods that require a bit more sophistication of both teacher and students, and some ideas that will not really be effective until the next generation of the Web.

Getting Started

The simplest way to begin using the Internet in your class is to put your e-mail address on your syllabus. All of the advantages of e-mail apply to your interactions with students. No phone tag, no garbled messages, you control when you respond. Especially for large classes and for shy students, e-mail allows you to communicate more effectively.

While e-mail is generally one-to-one communication, you can conduct class discussions outside of class with an electronic conference or bulletin board. An electronic conference handles the distribution of e-mail so that any message sent to the conference is redistributed to all the participants. You can use an electronic conference to remind students about upcoming tests or assignments, television specials or lectures that are relevant to the course. You can also use them to make sure that everyone gets an answer to a question that someone asked after class or by e-mail. The conference also is a way of encouraging students to talk to one another about the course. A bulletin board works similarly to a conference, but the messages are not delivered automatically to each student's e-mail account. Conferences and bulletin boards assume that all of your students have access to the campus computer network and have e-mail accounts, but that is increasingly common on campuses today. You will need to talk to someone at your college or university who is responsible for the campus network to find out how to create a bulletin board or electronic conference at your institution. Once you have created the conference or bulletin board, you will be able to re-use it from one semester to the next. You can use the conference or bulletin board to facilitate communication informally or you can include participation as part of your evaluation of each student at the end of the course.

You don't necessarily have to create your own conference or bulletin board. There are many that have been already established around broad and narrow topics. There are general archeology lists such as ARCH-L,[1] regional lists such as AZTLAN-L and topical lists such as HISTARCH. In addition, there are many USENET bulletin boards including sci.archaeology.mesoamerica, sci.archaeology, sci.anthropology.paleo and talk.origins. Two good sources of information about mailing lists are Anthropology Resources on the Internet[2] (formerly by Allen Lutins and now maintained by Bernard Clist) and the National Center for Preservation Technology and Training's Preservation Internet Resources[3] (Figure 1).

If you are going to ask your students to subscribe to one of these conferences, you should provide them with some guidance regarding etiquette. Lists with established subscribers do not always respond diplomatically to requests like, "I need to know about

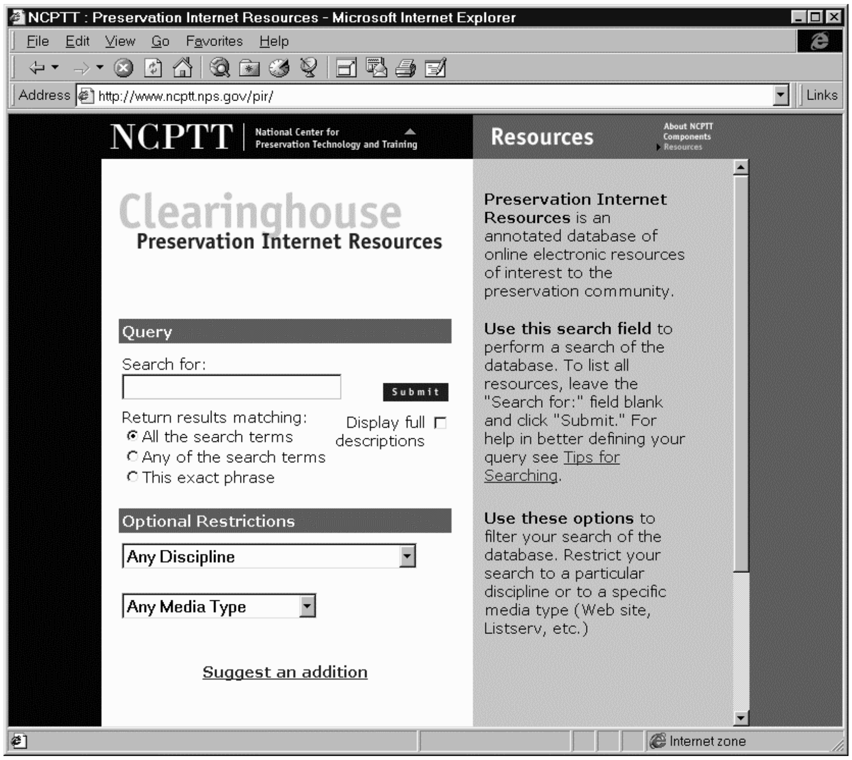


Figure 1. NCPTT's Preservation Internet Resources.

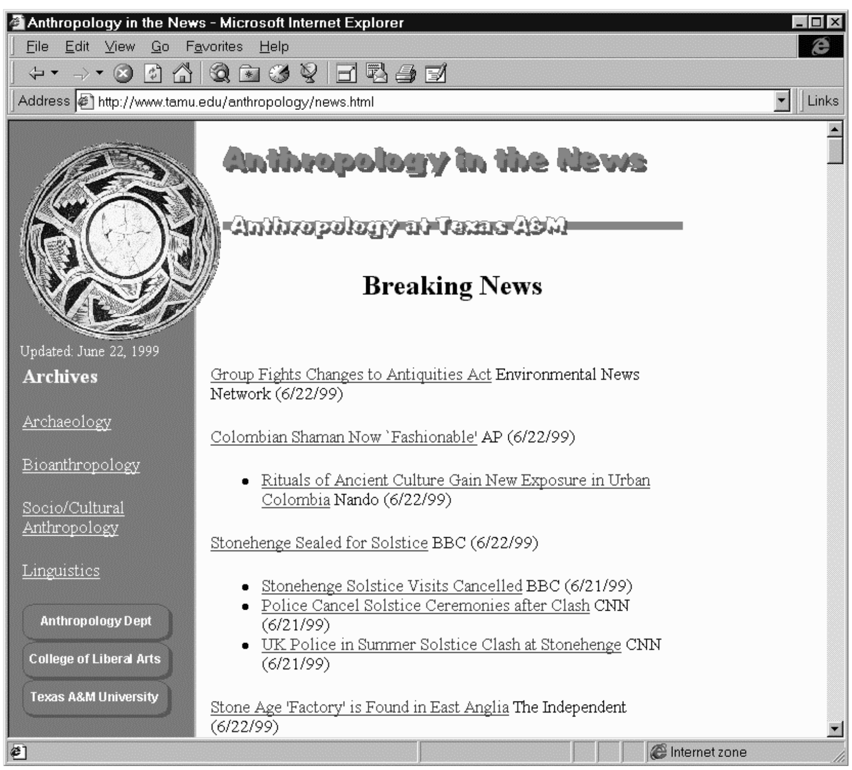


Figure 2. Anthropology in the News Web site. Reprinted with permission from the Anthropology Department, Texas A&M University.

some books on the Aztecs by tomorrow." On the other hand, good questions usually stimulate good answers and productive discussions that draw on the experience of archeologists around the world.

Creating Web Pages

If you are teaching a large lecture class, creating a class Web page will allow you to provide a variety of information to your students at virtually no cost to your department. I generally create a simple Web page for each class that contains a copy of the course syllabus, links to Web sites that are relevant to the class, copies of the transparencies that have my lecture outlines and any visual material that I use in class that is not copyrighted. I also put study guides for the tests on the class Web page and post test grades. For some classes I have developed collections of Web links that relate to class topics.

While creating visually engaging Web pages can be time-consuming, simple Web pages are created easily with software you already may have on your own computer. Current versions of Corel WordPerfect and Microsoft Word allow you to convert a file to html[4] format. While the results will not perfectly reproduce your original materials, it will probably suffice. You can improve the conversion by keeping several things in mind when you create a document that you plan to publish on the Web using a word processing program. Certain formatting codes that are common in word processing are missing from the current definitions of html. Tabs and indent codes are examples of formatting codes that do not exist in html. When making a syllabus with columns for dates and reading assignments, create a table in Word or WordPerfect instead of using tabs to create columns. Tables convert easily to Web documents. Lists also are defined in html so that creating a numbered list or a bulleted list is another way of indenting text.

If you want to work directly in html format, you probably will be better off using a Web page editor. Microsoft Internet Explorer and Netscape Communicator each come with Web page editors that will be adequate for most of your needs. The only missing step is getting your pages on the Web. Your university probably provides space on the university computers for your Web page, but you will have to find out how they want you to upload the information. This may involve another program or you may be able to retrieve and save your Web files from your Web page editor.

Web Resources

While setting up electronic conferences and class Web pages will help you reach your students more effectively, you also can use the material on the Web to stimulate their interest in archeology and anthropology. While you can attempt to load and display Web pages in class on a computer connected to a video projector, your students will not find this very stimulating. The Web is an interactive medium that works best when each student controls the pace and direction of the exploration. You will get better results by assigning activities to your students that are completed outside of class. Those activities should result in papers, class presentations or class discussions. You can create these activities yourself, but you should check with the publisher of your textbook since many of them are now establishing Web sites to support their texts.

The excitement of archeology and its relevance in the contemporary world are reflected in recent news stories that concern discoveries, great debates and controversies. All of the major news media now maintain Web sites that contain much of their printed or broadcast material. You can easily find links to news items relating to anthropology and archeology at Anthropology in the News,[5] a site that I started about two years ago to replace a bulletin board full of news clippings that I kept outside my office (Figure 2). The news items on the site can be used to stimulate discussion in class or on your class mailing list.

If you have spent any time surfing the Web, you know that archeology is well represented. If you don't know what is on the Web for archeologists, you should visit some of the Web sites that index other Web sites. For archeologists some good starting points are ARCHNET[6] — the grandfather of indexes to archeology sites, but no longer up-to-date — the World Wide Web Virtual Library for Anthropology,[7] Yahoo's Anthropology and Archaeology[8] section, Kris Hirst's Archaeology[9] page for About.com, and Anita Cohen-Williams' Archaeology on the Net.[10]

A simple way to begin incorporating information available on the Web is to provide links to Web pages that provide additional details to material covered in your lectures, in the text or in the documentary videos that you are using. The sites can be sources of more up-to-date statistical information or current events that relate to people, societies or other topics covered in the course. While these links are useful to students, you often will find that they do not use the material unless you are specific regarding how the material is to be used. The sites can provide a basis for classroom discussions or your students might use them to find ideas for research papers or reaction papers. Students also could be asked to review and critique the sites in papers or classroom presentations. You can ask for written reviews or critiques of the sites or can include questions about the sites on your tests. If you don't yet have a collection of relevant links ask students to find Web sites that relate to topics in the course. If your class is relatively small, you can make the creation of the course Web site a project for the entire class.

The various activities that you can organize around the Web fall roughly into four categories. First, you can assign a Web site or page as you would a reading assignment and ask students to learn the material presented there. Second, you can assign a Web site and ask students to critically evaluate the logic and evidence cited. Third, you can ask students to find information on the Web, at either a Web site addressing a topic or a Web page containing specific information. Finally students can use the Web to master course materials through interactive quizzing.

Supplementing Traditional Course Materials

The simplest Web assignment is one in which you ask students to read a specific document. This activity uses the Web as a kind of 24-hour reserve room and is a good way for students to begin to become comfortable with the Web. The variety of articles available on the Web is limited but it is growing. *Scientific American*, *American Scientist*, and other magazines put one or two articles from each issue online. In other cases, authors put copies of published articles on the Web, or put unpublished or in-progress work on the Web. As examples, the following articles are available:

- "The African Emergence and Early Asian Dispersals of the Genus *Homo*"[11] from *American Scientist* by Roy Larick and Russell L. Ciochon
- "An Evaluation of Chaco Anasazi Roadways"[12] — a paper presented at the 1996 SAA Meetings by John Kantner
- "The Origin of the Human Capacity"[13] — the Sixty-Eighth James Arthur Lecture on the Evolution of the Human Brain, given at the American Museum of Natural History by Ian Tattersall (Figure 3)
- "The Viking Longship"[14] in *Scientific American* by John R. Hale
- "Getting Their Hands Dirty: Archaeologists and the Looting Trade"[15] in *Lingua Franca* by John Dorfman
- "Transitions in Prehistory"[16] in *Science* by Tim Appenzeller, Daniel Clery and Elizabeth Culotta

Your university may have electronic versions of scholarly journals that your students can access. Electronic versions of the Annual Review series and Academic Press journals are available now and others eventually will be available.

In addition to assigning an article on the Web, you can assign audio or video clips. The availability of these also is limited but growing. National Public Radio maintains an archive of programs and interviews that can be played with a RealAudio plugin. National Geographic Society,[17] PBS,[18] and the Discovery Channel[19] also have audio programs available. Some examples include the following programs –

- "First Humans in the Americas"[20] with Michael Waters, Thomas Dillehay, Dena Dincauze and Roger Powers, February 28, 1997 *Talk of the Nation*
- "Human Origins"[21] with Donald Johanson, May 9, 1997 *Talk of the Nation*
- "Early Human Ancestors"[22] with Antonio Rosas, May 30, 1997 *Talk of the Nation*
- Interview with Mark Lehner[23] from "Pyramids, The Inside Story"[24] from *NOVA*.
- "Discovering the Maya"[25] by George Stuart, National Geographic Society, April 14, 1999 (Figure 4)
- "Demille Dig"[26] Renee Montagne reports on an archeological excavation along the coast of California to recover Cecil B. Demille's set for "The Ten Commandments" on *Sounds Like Science* for National Public Radio, April 3, 1999

Video still is relatively rare on the Web because the storage requirements are so great and the quality is still low, but clips of recent news stories are available on many different news sites including CNN[27] and ABC.[28]

Articles, audio clips, and video clips are relatively easy to incorporate into your course since they are linear media. All students proceed from beginning to end in the same sequence so it is relatively easy to define what they should learn in the process. One of the advantages of the Web is that multimedia presentations need not be linear, which means that visitors to a site may all begin at the same place, but then diverge into different directions. Archeology has a relatively large number of multimedia sites that use a combination of text, images, sound, video or virtual modeling to describe an archeological site or to discuss an archeological topic. Your students can get much out of these sites, but you will have to be specific regarding how much of the site they need to visit.

Web sites that focus on particular archeological sites have a number of advantages over printed versions. Publication to the Web is fast and inexpensive. Web treatments of archeological sites have even been developed simultaneously with excavation. Color images cost no more to reproduce than line drawings. They cost students nothing to use. On the other hand, they usually go through fewer stages of review and, once created, they can linger on the Web after their information has become obsolete. While the sites can be a valuable complement to

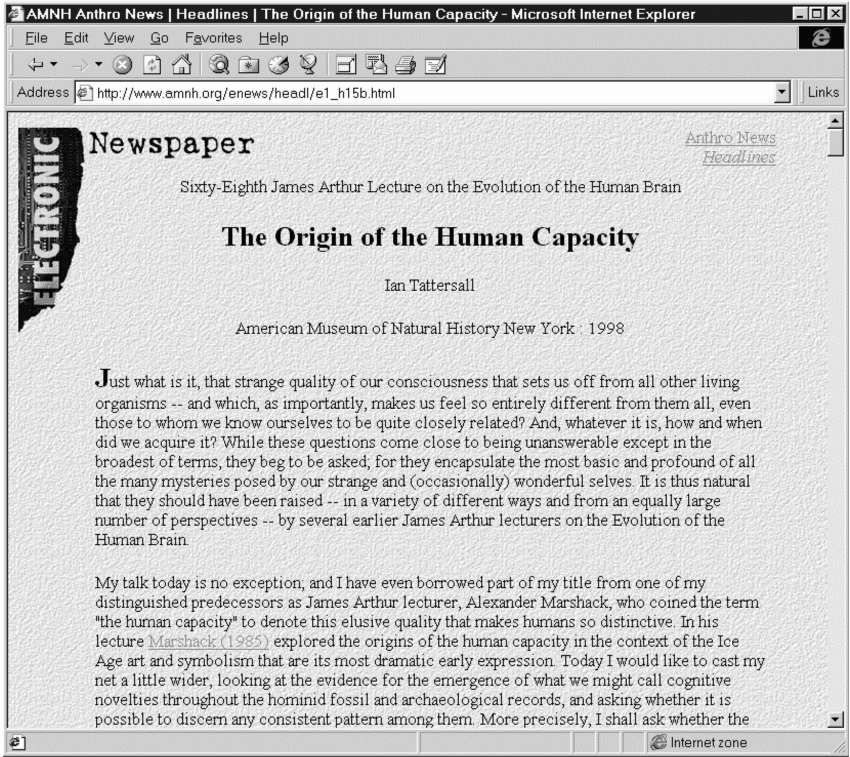


Figure 3. “Evolution of the Human Brain” lecture by Ian Tattersall. Reprinted with permission from Ian Tattersall and the American Museum of Natural History.

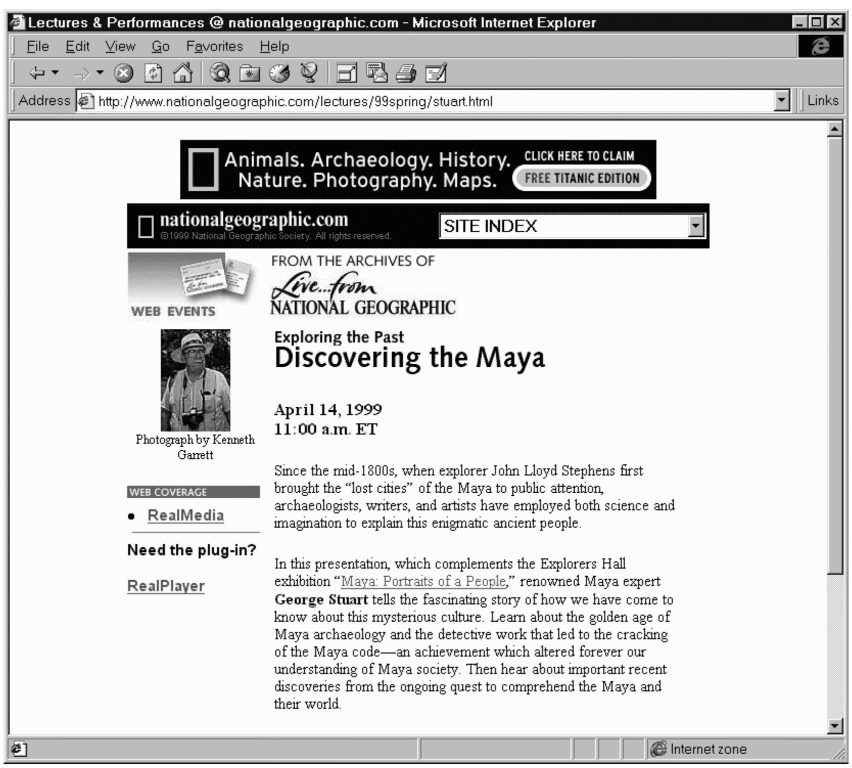


Figure 4. “Discovering the Maya” lecture by George Stuart. Reprinted with permission from the National Geographic Society.

teaching about archeology, you will need to exercise quality control by selecting only sites that are accurate, current and present archeology as more than the collection and illustration of interesting artifacts. There are a number of good sites available, such as:

- The Jamestown Rediscovery Site[29] by the Jamestown Rediscovery Project
- Keatley Creek,[30] Charlie Lake,[31] and Namu[32] at Simon Fraser University
- Five Points, New York City[33] by Rebecca Yamin
- La Grotte de Lascaux[34] by the France's Ministry of Culture
- The Ceren Web Resource, Joya del Ceren, Archaeological site, El Salvador[35] at the University of Colorado
- Çatalhöyük[36] at Cambridge University (Figure 5)

Providing less detail about sites, but fun to explore are a variety of three-dimensional reconstructions of archeological sites including:

- The 3-D Reconstruction of Chetro Ketl Great Kiva[37] by John Kantner
- Tikal[38] by Studio360
- Tenochtitlan[39] by Dell Maxwell
- Ancient Hrappa[40] by Wayne Belcher
- Tunnels of the Western Wall of the Temple Mount in Jerusalem[41] by Aish HaTorah
- Virtual Palenque[42] by Qvision (Figure 6)

Other Web sites address a particular topic. Many of these have been designed around documentary programs so that they provide a nice complement to the program if you are using the video in class. Most of them will also stand on their own and allow students to explore a topic on their own in more detail than their text or classroom presentations.

- "Andes Expedition: Search for Inca Secrets"[43] by National Geographic Society
- "Secrets of Easter Island"[44] by *NOVA*
- "In the Footsteps of Alexander the Great"[45] — Michael Wood's series on PBS
- "Collapse: Why Do Civilizations Fail?"[46] from the *Out of the Past* series by Annenberg/CPB (Figure 7)
- Treasures of the Sunken City[47] by *NOVA*

Developing Critical Skills

One of the biggest concerns about material on the Web is how reliable it is. You can use archeology and the Web to help students develop their critical skills when it comes to evaluating claims made at Web sites. As you would expect, there are Web sites that talk about how to evaluate Web sites critically. For example, Internet Detective[48] is an interactive tutorial in how to evaluate the quality of Web resources (Figure 8). Other good pages are Critical Thinking Resources[49] at Longview Community College and A Student's Guide to WWW Research: Web Searching, Web Page Evaluation, and Research Strategies[50] by Craig Branham at St. Louis University. Once you have discussed evaluating Web sites, you can provide a link to a Web site and ask your students to evaluate its credibility. Alternatively you could ask students to compare two Web sites such as a fantastic archeology site and a site critiquing those claims. The Web is not the home to more outlandish claims about archeology than you will find on television or at the newsstand, but the Web makes it easy to place claim and counterclaim side-by-side.

Building Research Skills

As your students become familiar with the Web, you can assign the task of finding types of sites on the Web. You may ask students to find one or more Web sites that are designed for a specific audience. In doing this, they will become more experienced at using the variety of search engines that are available. Some engines, (such as Yahoo!), are better at finding Web sites, as opposed to specific information on specific pages within a site. There is no correct answer but students learn how to find sites on the Web and get the flexibility of seeking sites related to their individual interests. In a similar activity you may ask students to find specific data on the Web. This activity is slightly more challenging since students must evaluate the quality of alternate sources. In some cases there may be more than one correct answer so that evaluation of this activity should focus on the process of locating and evaluating the information more than the specific answer.

Asking your students to learn about an issue on the Web is more challenging. Your question may be posed more broadly and the issue can involve strong proponents for opposite positions. Students may be asked to locate two or more competing positions and analyze the issue in terms of the claims by each side. This activity involves more skill in searching the Web for information and sifting through numerous possible Web sites for those that are relevant. The activity also involves a critical evaluation of two or more positions. This activity could be used as the basis for an essay question on an exam or as a springboard for class discussion. Combined with library research, this activity could be the basis for a research paper.

While the Web provides a great deal of information to assist students in learning about anthropology and the world around them, the Web can also tempt them to bypass the library and limit themselves to Internet resources. You, or a library representative, should talk to students about how students can use the Web to improve research skills with books and journals

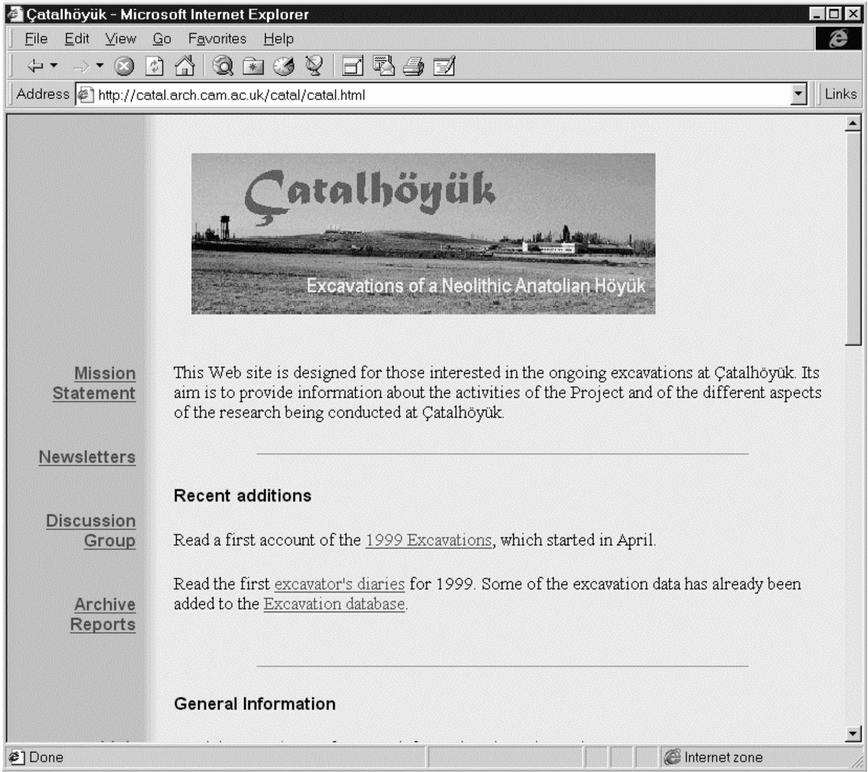


Figure 5. Çatalhöyük Web site. Reprinted with permission from Cambridge University and the Çatalhöyük Project, Ian Hodder, Director.

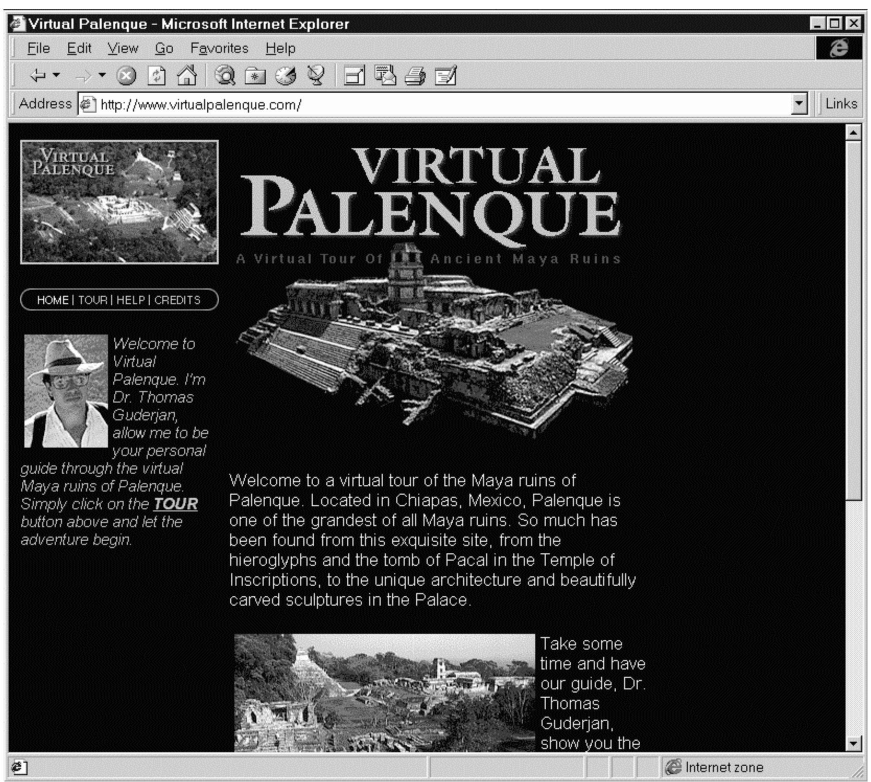


Figure 6. Virtual Palenque Web site. Reprinted with permission from QVision.

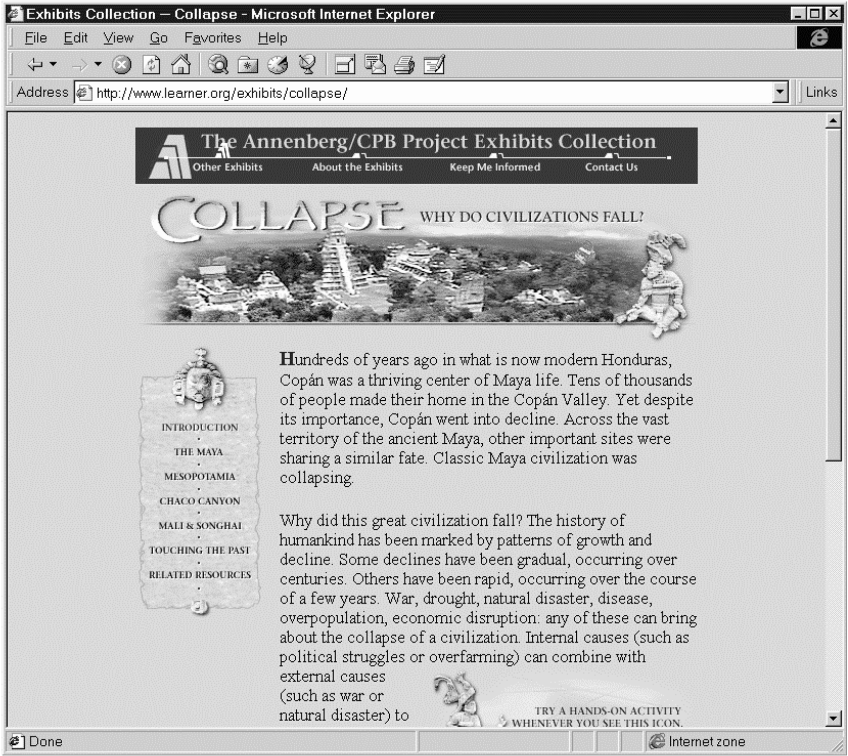


Figure 7. Collapse: Why Do Civilizations Fail? Web site. Reprinted with permission from The Annenberg/CPB Project.

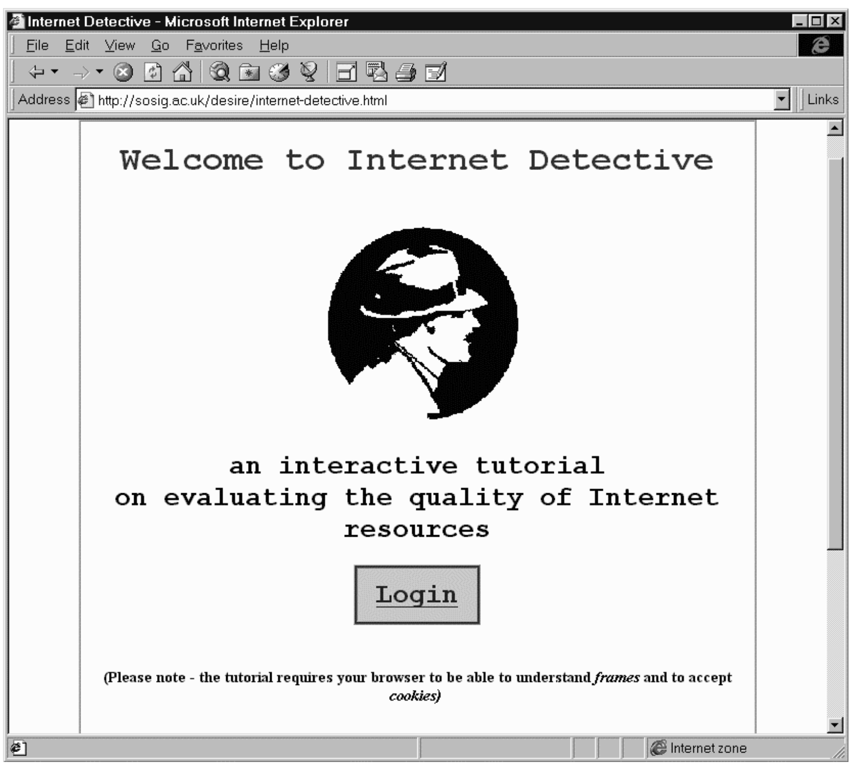


Figure 8. Internet Detective Web site. Reprinted with permission from the University of Bristol.

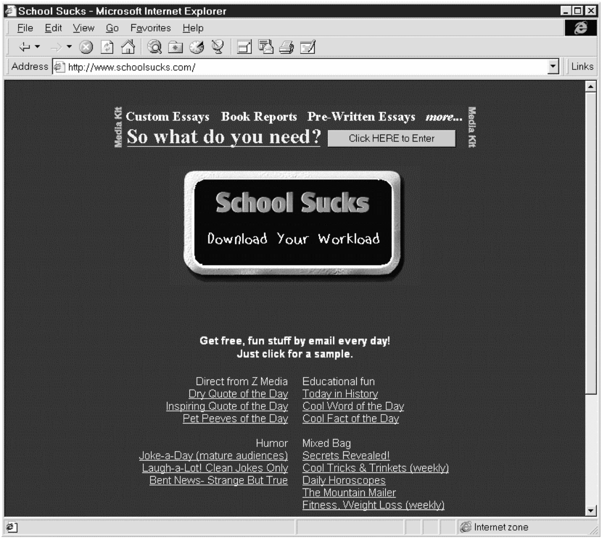


Figure 9. School Sucks Web site. Reprinted with permission.

in the library. Help students find out how to access the university library catalog online and let them know which journal indices are available online at your institution. You should also talk about plagiarism and the Web. You will get some useful ideas on Gregory Senechal's Instructor's Guide to Internet Plagiarism[51] site and from Tom Rocklin's article "Downloadable Term Papers: What's a Prof to Do?"[52]

You should be aware of the large number of services on the Web that provide term papers to students. A few sites, such as School Sucks[53] (Figure 9) provide papers for free to students (and their instructors). Other services charge for papers. The cost is usually about $5 to $10 per page for pre-written papers and more for custom-written papers. You can find sites like these at Yahoo's Research and Term Papers[54] section or by using a search engine for term papers or research papers. There are now a few sites that claim to evaluate papers for plagiarism by comparing them to a database of papers, but it is very unlikely that the database is really complete or could possibly include "custom" papers. Tom Rocklin suggests focusing on the process of writing a term paper. Require students to select their paper topic early in the term. Require a bibliography and an outline of the paper. Have students give a brief presentation on their paper with time for questions from other students or the instructor. While none of these guarantee that a student will not take the easy way out, they make it somewhat more difficult than if the paper is simply announced at the beginning of the term and collected at the end of the term.

Interactive Quizzing and Tutorials

Many major publishers are building Web sites for their texts and increasingly offer various interactive activities such as quizzing. You should try these activities yourself before deciding what to require or recommend. Since there is no security for online quizzing you probably will not want to use the quiz scores directly. The quizzes may, however, be a useful study tool in helping students to master the material.

Conclusions

The Web has the potential to enhance and improve the quality of teaching about archeology. By providing students with direct access to current information and to diverse claims and counterclaims, the Web helps us to communicate the process that archeologists employ to understand the past. In that sense, the Web does not replace texts or the library, but provides additional means to help students see how we come to conclusions about the past.

Notes

1. <listserv.tamu.edu/archives/arch-l.html>

2. <home.worldnet.fr/clist/Anthro/index.html>

3. <www.ncptt.nps.gov/pir/>

4. <dir.yahoo.com/Computers_and_Internet/Internet/World_Wide_Web/Information_and_Documentation/Beginner_s_Guides/Beginner_s_HTML/>

5. <www.tamu.edu/anthropology/news.html>

6. <archnet.uconn.edu/>

7. <www.anthrotech.com/resources/>

8. <www.yahoo.com/Social_Science/Anthropology_and_Archaeology/>

9. <archaeology.about.com/>

10. <www.serve.com/archaeology/>

11. <www.amsci.org/amsci/articles/96articles/Larick.html>

12. <sipapu.ucsb.edu/roads/index.html>

13. <www.amnh.org/enews/headl/e1_h15b.html>

14. <www.sciam.com/1998/0298issue/0298hale.html>

15. <www.linguafranca.com/9805/dorfman.html>

16. <www.sciencemag.org/cgi/content/full/282/5393/1441>

17. <www.nationalgeographic.com/>

18. <www.pbs.org/>

19. <www.discovery.com/past/sciencelive/morearcheology.html>

20. <www.npr.org/ramarchives/nf7f2801-6.ram>

21. <www.npr.org/ramfiles/970509.totn.02.ram>

22. <www.npr.org/ramfiles/970530.totn.01.ram>

23. <www.pbs.org/wgbh/nova/pyramid/excavation/lehner.html>

24. <www.pbs.org/wgbh/nova/pyramid/>

25. <www.nationalgeographic.com/lectures/99spring/stuart.html>

26. <www.npr.org/ramfiles/sls/19990403.sls.07.ram>

27. <www.cnn.com/>

28. <abcnews.go.com/>

29. <www.apva.org/>

30. <www.sfu.ca/archaeology/museum/bc/keat_src/k0000001.htm>

31. <www.sfu.ca/archaeology/museum/bc/clc_src/cl000001.htm>

32. <www.sfu.ca/archaeology/museum/bc/namu_src/index.htm>

33. <r2.gsa.gov/fivept/fphome.htm>

34. <www.culture.fr/culture/arcnat/lascaux/fr/index.html>

35. <ceren.colorado.edu/>

36. <catal.arch.cam.ac.uk/catal/catal.html>

37. <sipapu.ucsb.edu/html/kiva.html>

38. <www.destination360.com/tikal.htm>

39. <www.cse.cuhk.edu.hk/~csc5460/mirror/handbook/>

40. <www.harappa.com/3D/index.html>

41. <www.aish.edu/tours/tunnel/index.html>

42. <www.qvision.com/palenque/>

43. <www.nationalgeographic.com/andes/index.html>

44. <www.pbs.org/wgbh/nova/easter/>

45. <www.pbs.org/mpt/alexander/>

46. <www.learner.org/exhibits/collapse/>

47. <www.pbs.org/wgbh/nova/sunken/>

48. <sosig.ac.uk/desire/internet-detective.html>

49. <www.kcmetro.cc.mo.us/longview/ctac/toc.htm>

50. <www.slu.edu/departments/english/research/>

51. <www.ab.org/gregg/>

52. <www.uiowa.edu/~centeach/newsletter/online/term-paper-download.shtml>

53. <www.schoolsucks.com/>

54. <dir.yahoo.com/Business_and_Economy/Companies/Communications_and_Media_Services/Writing_and_Editing/Research_and_Term_Papers/>

A New Way to Publish: Journal Databases Evolve on the World Wide Web

HUGH W. JARVIS

As the impact of the World Wide Web is deepening, journal publishing is evolving to suit this new niche. While paper periodicals continue to be produced, new forms of publications are emerging. An exponentially increasing online readership, significant savings in production and distribution costs, much faster dissemination rates, and the potential for interactive and enhanced documents are the lure. An unforeseen side effect is that there is really no longer a need for bundled "periodicals". A number of anthropological publications are helping to lead this revolution.

Journal publishing is a complex business that includes the challenge of channeling the work of scholars into a form useful to their colleagues. The advent of the World Wide Web as a publishing medium adds yet another dimension to this complex business, solving some problems while creating new complications and twists. An evolutionary transformation of journal publishing has begun as journals migrate online and change in the process.

Is the Web a Reality for Publishing?

Initially, let's look at the current potential of the Web and the Internet as a publishing medium.

Internet Use

The first question one might ask is how many people actually use the Internet? Unfortunately, accurate Internet use is almost impossible to measure.

One way to gauge at least potential Internet use is through Census data (see Table 1), which shows acceleration in the presence and use of computers throughout the United States. The number of households with computers has doubled every five years. Additionally, the number of people who may not actually own a computer, but use one in the course of their regular lives — be it at school, work, home — has risen about fifty percent every five years (see Table1).

Unfortunately there is no corresponding progression of data for Internet use. For 1993, the only year from which data is available, 10 million people (4% of the population) spent at least some time reading e-mail at home.[1] Presumably this number is much larger now, but we will not know for sure until the upcoming 2000 census has been completed.

Despite the recent attempts by Microsoft and Intel to begin tagging individual computers and stories that appear in the popular press, tracking or analyzing online activity is virtually impossible. Instead, most knowledgeable analysts use more robust measurements. One approach is to count the base number of host servers — the means by which people connect to the Internet — and then extrapolate anywhere from one to twenty users per host. While this count may not provide an exact number of Internet users, it does closely reflect changes in the level of demand for Internet access.

Fantastically, the number of hosts has been doubling about every year (see Figure 1), from only four in the comparatively early days of 1970, to a staggering 43 million by January 1999. This exponential increase suggests anywhere from 43 to 430 million users worldwide, and those figures already are several months old.[2]

Another way to indirectly measure users is by examining the level of Internet traffic itself. The data in Table 2 show that the amount of digital data moving around the Internet has been rising at an increasing rate, from just 200 megabytes per month in 1980, to 100 million megabytes per month in 1996. This flow is doubling almost every hundred days, and the growth rate is unlikely to level off soon.[3]

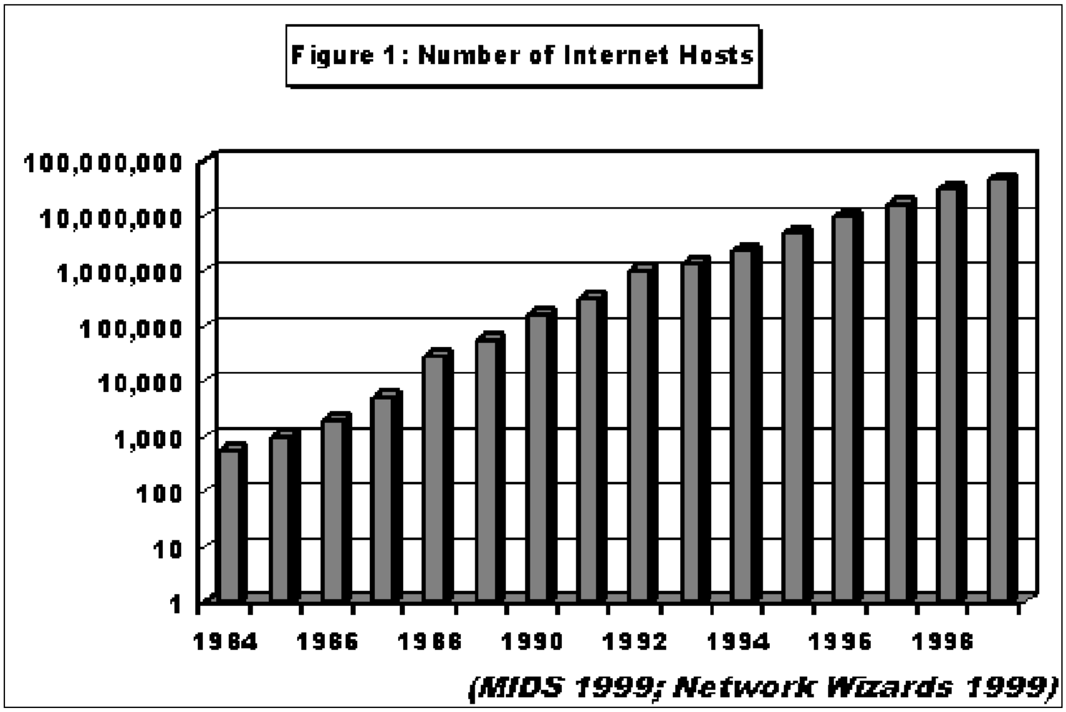
Table 1: US Census Data – Computer Use

	1984	1989	1993
Households With Computers	7 million (8%*)	14 million (15%*)	23 million (23%*)
People Using Computers	47 million (21%*)	75 million (32%*)	100 million (41%*)

** percent of total population (US Census 1993)*

Web commercial productivity

In addition to reading e-mail and downloading pictures of archaeological site excavations, Internet users are so eager for Web information, that they are creating their own sites at an astronomical rate (see Figure 2). The number of Web sites, another difficult phenomenon to count, has grown from fifty in 1992 to an estimated 1.2 million in 1997. And not only are these people avid Web readers, they also are willing to spend hard cash. Electronic commerce has grown 400 percent per year, from $10 million in 1996 to an estimated $240 million in 1998, with sales projected to reach the tens of billions or even a trillion dollars by 2002.[4]



Journals Follow Suit

Publishers have not ignored these developments, and the number of electronic journals and newsletters is growing very rapidly, as shown by Table 3. Data have been collected only since 1991, when there were 110 such periodicals. Since then, the numbers have grown to 3,414 in 1997, of which 1,049 are peer reviewed. Of these, 28 percent are focused on social science topics. Eighty-six percent are scholarly in nature — arts and humanities, life sciences, physical sciences, technology, and social sciences — and 14 percent are categorized as recreation and general interest.[5] By contrast, there were an estimated 7,000 print journals in 1995 and about 14,000 journals of all types in 1999.[6]

Recognize that these numbers reflect more than just the birth of online journals. In the last few years, publishers have begun to put some or even all of their publications online, in addition to print versions. Indeed, the American Chemical Society now has full text equivalents online for all of its journals[7] and commercial publishers like Reed-Elsevier (over 1,200 journals), Springer (360 journals), Academic Press (174 journals) are following suit.[8]

Discussion

In summary, there is an exponentially growing market of readers willing to spend increasingly large sums of money online and desperate to read information, and publishers are shifting to meet this demand by migrating online.

Table 2: Monthly Internet Traffic in Megabytes

1980	1982	1984	1986	1988	1990	1992	1994	1996
200	500	2,000	10,000	100,000	1,000,000	5,000,000	12,000,000	100,000,000

(MIDS 1999; Network Wizards 1999)

Why publish a journal online?

Still, one might ask, why would traditional publishers want to spend time and money developing new formats for the electronic medium? Surely they are quite successful already. Why would they or their subscribers want a change?

Speed of dissemination

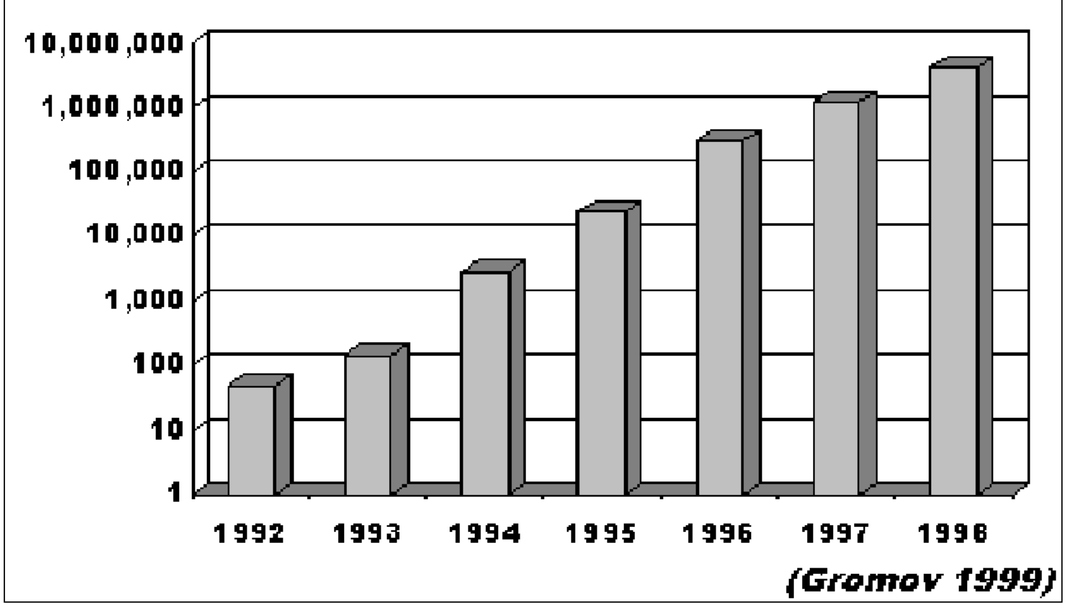
With online publication, the production cycle is significantly shorter. Issues can be published within hours of their clearing the editing and layout stages, months or even years less time than it would take for them to complete the printing, binding and distribution process.[9] If the individual article becomes the unit of production, rather than the whole issue, this process is shortened even further.

Enhanced / value-added features

Online publications can offer a wide range of features that are simply not possible for their print relatives.[10] Examples include:

- no article size limits;
- no limit to the number of graphs, tables and other figures;
- internal hyperlinks linking text sections and to references, tables, and appendices;
- external hyperlinks connecting to the authors' e-mail and to online versions of cited or relevant references and Web sites;
- links to expanded data sets, including entire external databases with search interfaces;
- complex multimedia such as picture galleries, three-dimensional images, video clips and audio files;
- attachments such as reader and editorial comments;
- embedded software so that readers can test their own data; and
- "living articles", such as dynamic — or even interactive — ongoing experiments.

Figure 2: Number of Web Sites



Reduced subscription / publishing costs

It is possible to see a direct economic benefit from shifting a journal online. Naturally, the eventual real savings will depend on the nature of each publication. Commercial publications aim for profit, while scholarly societies tend to be non-profit or even to publish at a loss. Scientific journals carry much more complex graphical information than those in the humanities. The actual baseline costs of publications vary, with some printed and bound quite cheaply, while others are produced much more expensively. Some publications are mostly text, while others include expensive images, equations, or figures, and have associated costly layout and production concerns.

Tenopir and King[11] discuss the costs of publishing journals, primarily based on scientific journals. "First copy" costs, such as review administration, editing, illustration preparation and layout, can reach an annual total of $200,000. In addition, marketing, subscriber maintenance, amortization of startup costs and overhead can run another $200,000 per year. The manufacturing process, the cost of paper, printing and binding, plus the cost of distribution averages $30 per subscriber. While the first two costs are somewhat independent of the medium, online publications have no physical manufacturing or distribution costs. These savings can be retained by the publisher as increased profit or passed along to the subscribers. Indeed, it would only be fair to reduce the prices of online editions, since a fraction of what the publishers save in terms of printing and distribution is actually passed along to the end-user for Internet connection fees and local printing costs. Prices also should equalize for all individuals as the higher distribution costs for international subscribers no longer apply.[12]

Journals have been caught in a pricing spiral. In an attempt to increase revenue, publishers have raised subscription fees. Personal subscription prices rose 85 percent in real dollars per decade over the last twenty years. In response, personal subscriptions have fallen from an average of 4.2 subscriptions per university scholar in 1977 to only 3.9 in to 1993.[13] Over the same period, non-university subscription levels have fallen even further, from 6.2 to 2.6.[14] To compensate, publishers raised prices even more, and the vicious trend has continued. Readers have turned to libraries to fulfill their journal needs but, except for popular journals, many items not available locally must be sought through expensive interlibrary loan and document delivery services. The net effect is no real savings to users, since they ultimately pay for rising library subscription costs and special delivery fees. More importantly, users have been subjected to a sharp rise in inconvenience and access delays, while publishers have suffered large revenue loss.

Discussion

Online journals offer a wide range of features not available in print formats. They also can be produced much faster and cheaper than traditional publications. Whether this saving will be used to restore falling profits or passed on to subscribers is not yet apparent.

How is Publishing Evolving to Meet This Opportunity?

As noted above, publishers have not failed to take advantage of the Internet as a medium. But the transition is much more extensive than this.

Publishing is Going Online

Do not be fooled by the continued presence of print versions. Robert Bovenschulte, Director of the American Chemical Society Publications Division, feels that electronic journals will completely supersede print journals within ten years, and Peter Boyce, an associate at the American Astronomical Society, feels this will occur in just three years.[15] While some print publications almost certainly will continue to be produced, there is a significant migration occurring to the Web. While this transition may be slower for the social sciences and humanities than the hard sciences, this change is so credible that many libraries are discontinuing print subscriptions, while some, such as the Technical Knowledge Center and Library in Lyngby, Denmark have phased out print publications altogether.[16]

Publishing is Shifting Away From Bundled Articles

A second phenomenon also is occurring. For example, all American Chemical Society journal articles are published online as soon as they complete the editorial process. They call this "ASAP" — "As Soon As Publishable" — and boast this allows scholars to access them as much as eleven weeks earlier than they will appear in print.[17] The National Aeronautics and Space Administration has funded a project, the Astrophysics Data System, that includes an abstracting service with over 1.2 million items accessible in an online database.[18] A related project is the Los Alamos National Laboratory e-Print Archive.[19] Started in 1991, this online database houses preprints of articles in physics, mathematics, neuroscience and computer science that are submitted directly by their authors. These services have become the primary means of communication for scholars in these disciplines. Plans to develop a similar project are underway by biomedical scientists.[20]

Print periodicals typically suffer scheduling and manufacturing constraints, related to bundling individual articles into journal volumes. On the Web

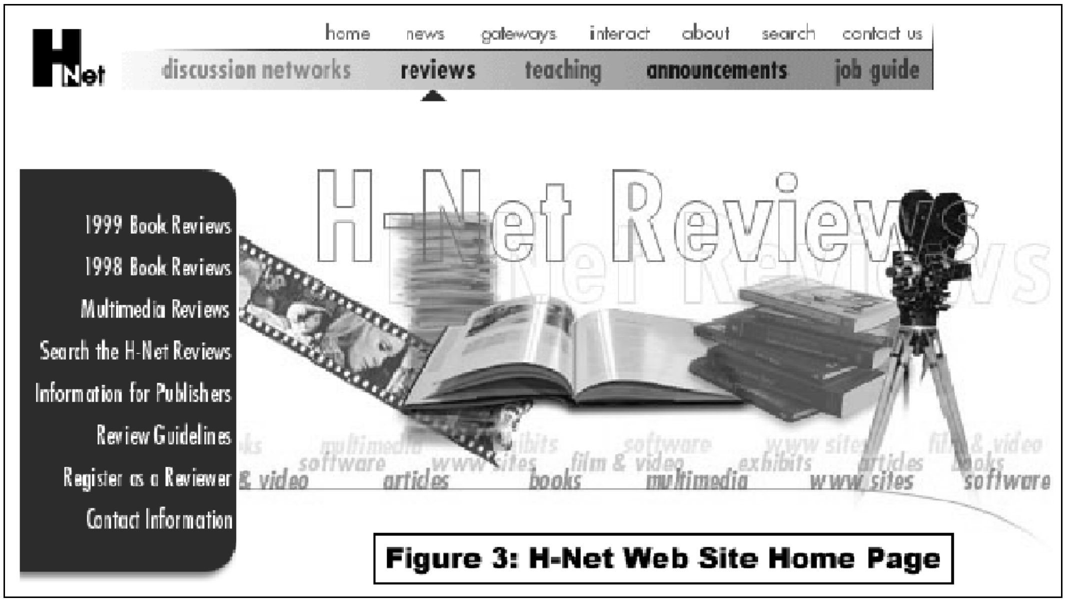


Figure 3: H-Net Web Site Home Page

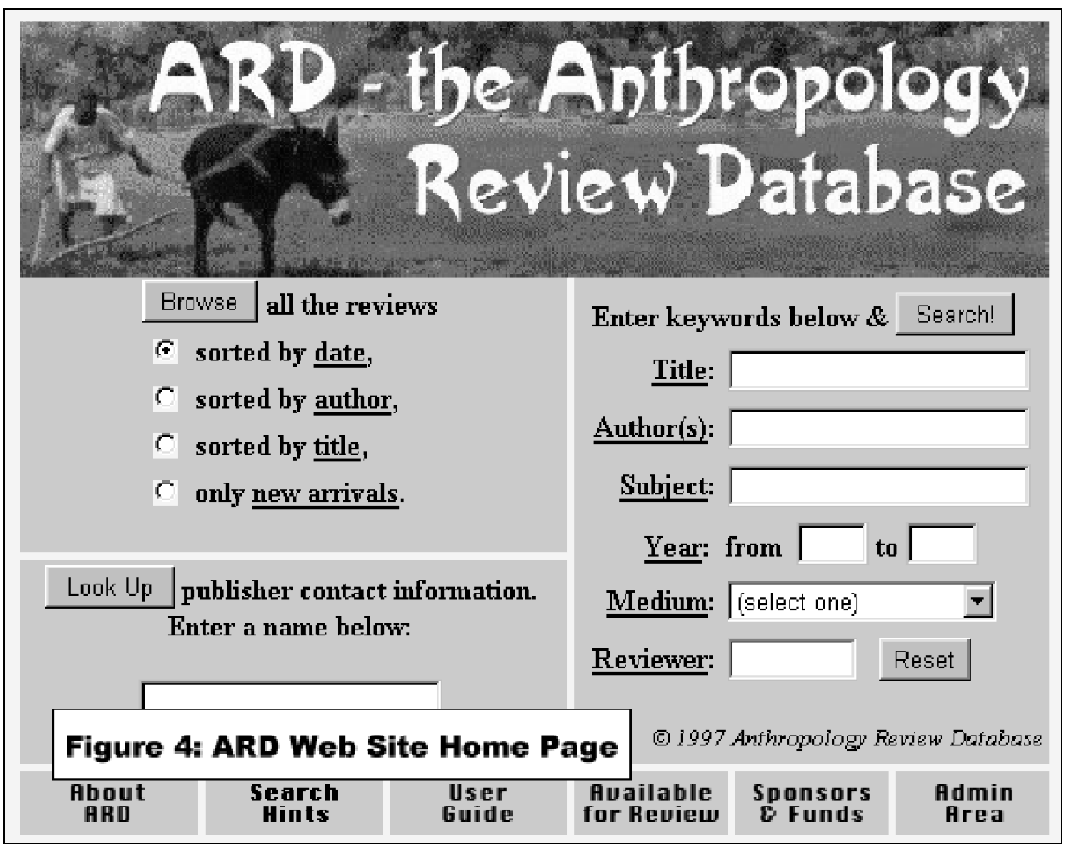


Figure 4: ARD Web Site Home Page

the need for bundling disappears. A handful of publishers in the humanities and social sciences have taken the publishing process to its logical next step, producing dynamic databases of articles instead of periodicals. As examples –

- **Reviews in History** is published by the Institute of Historical Research at the University of London and began publication in 1996.[21] They publish scholarly reviews of two to three thousand words covering books on European and UK history that are available on their site.[22]
- **AnthroGlobe** is an international project located in Vancouver, Canada.[23] AnthroGlobe is intended to provide a friendly platform for authors to post drafts or completed works. The items are not refereed. Readers' comments are encouraged and the articles are expected to evolve on-site.
- **H-Net Reviews** (Figure 3) is a project based at Michigan State University.[24] H-Net lies at the heart of a large number of e-mail discussion lists that focus on many aspects of history. Once a review is posted to a list, it is housed in the main database.
- The **Anthropology Review Database** (Figure 4) is a project at the University at Buffalo Department of Anthropology.[25] Launched in 1997, and run solely by an international network of volunteers, ARD currently has about 250 items in its online database, including both its own refereed reviews and links to reviews located in other publications.
- The **Bryn Mawr Electronic Resources Review** is published by the Bryn Mawr College - Center for the Study of Architecture. This new project is producing reviews of electronic resources, including CDs and Web sites. The reviews are available in an online database[26] and also are distributed through e-mail lists as they are produced.

While most of these projects have chosen to focus on publishing reviews, the model should work for more traditional journal articles as well.

Discussion

Journal publishers are shifting to online formats, while some, particularly in the hard sciences, predict a complete digital migration within three to ten years. Additionally, some online publications no longer use a bundled format of individual articles in issues and have begun to publish items individually as they are completed.

Summary

The Web is a dynamic medium with exponential growth and tremendous potential for publishers. Online publications have numerous advantages over their print relatives, not the least of which are the potential for savings in production and distribution costs. Accordingly, many journals have begun to migrate online and there is strong indication that this may represent the beginning of the end for print publications. A final transformation may be that journals will no longer be periodicals, per se, and instead articles will be published individually.

Notes

1. *Table A. Level of Access and Use of Computers: 1984, 1989, and 1993*. US Census. <www.census.gov/population/socdemo/computer/report93/compusea.txt>.

2. *Internet Hosts Worldwide*. Matrix Information and Directory Services, Inc. (MIDS). <www.mids.org/growth/internet/html/hosts.html>
Internet Domain Survey. Network Wizards. January 1999. <www.nw.com/zone/WWW/report.html>.

3. Gregory Gromov. *History of Internet and WWW: The Roads and Crossroads of Internet History*. Internet Valley, Inc. 1999. <www.internetvalley.com/intvalstat.html>.

4. Ibid.

5. Dru Mogge, "Forward." *ARL Directory of Electronic Journals, Newsletters and Academic Discussion Lists*. 7th Edition. Washington: Association of Research Libraries. 1997. <www.arl.org:591/foreword.html>.

6. Sophie Wilkinson, "Electronic Publishing Takes Journals into a New Realm," *Chemical and Engineering News*. 76.20 (1998): 10.
"About the Gale Database of Publications and Broadcast Media." *Gale Database of Publications and Broadcast Media*. Detroit: Gale Research Inc. 1999. <galenet.gale.com/m/mcp/db/gdpbm/intro/about.html>.

7. Wilkinson, 12-13.

8. Declan Butler, "The Writing is on the Web for Science Journals in Print," *Nature* 397(1999): 195-200.

9. Wilkinson, 12.

10. Ibid.

11. Carol Tenopir and Donald W. King, "Setting the Record Straight on Journal Publishing: Myth vs. Reality," *Library Journal* 121.5 (1996): 32-35.

12. Ibid.
Wilkinson, 14.

13. Tenopir, 33.

14. Ibid.

15. Wilkinson, 12.

16. Butler, 195.

17. Wilkinson, 12.

18. <adswww.harvard.edu>.

19. <xxx.lanl.gov>.

20. Butler, 195.

21. <ihr.sas.ac.uk>.

22. Anne Shepherd, "Book reviews on the Net," *History Review* 28 (1997): 55-57.

23. <www.webzines-vancouver.bc.ca/ AnthroGlobe/>.

24. <www.h-net.msu.edu/reviews/>.

25. <wings.buffalo.edu/ARD>.

26. <csa.brynmawr.edu/bmerr>.

Where Have all the Data Gone? Issues in Web Site Design

MARY S. CARROLL
BART MARABLE

Publishing archaeological data and information on the World Wide Web presents a unique set of challenges. Along with technical issues such as data standards and file formats come the often-overlooked issues of Web site accessibility and design. Users may never get to the information available if they can't easily find their way around the site — or if the structure of the site cripples downloading capabilities. This paper will discuss the process of developing useful, accessible Web sites that deliver substantive information and will use the National Center for Preservation Technology and Training's Web site design project as an example.

Delivering archaeological information via the World Wide Web presents a unique set of challenges. Along with technical issues such as data standards and file formats come the often-overlooked and related issues of Web site accessibility and design. As the Web expands and technological capabilities grow, users expect more substantive content from the sites they visit and Web site managers attempt to deliver that content. But even though there may be a wealth of archeological information present on the Web,[1] archeologists searching for raw data, research reports, bibliographies and other resources may never get to the information available because they can't easily navigate the site.

Much has been written in the Web design world about how to develop cutting edge Web sites, both creatively and technologically.[2] In this paper we will briefly discuss the process of developing useful, accessible Web sites that deliver substantive information — and will use the National Center for Preservation Technology and Training's Web site design project as an example.

Even though the term "electronic publication" often is applied to the Web, well-designed Web sites are not publications — they are organized spaces that connect people with information and with each other. In fact, good Web site design has more in common with good exhibit design than with preparing a publication. Both Web sites and exhibits must entice their visitors, provide clear paths of exploration, keep visitors' attention and connect visitors with what they are looking for — from interactive experiences to in-depth informational resources.

Developing a well-designed Web site can be divided into five phases. Project planning — during which the site's objectives and information design are established — consists of the first two phases, definition and architecture. Project development, during which the site is designed and constructed, takes place in the latter three phases — design, implementation and integration. While jumping directly to the development phases may be tempting, the ultimate success of a Web site depends on careful planning prior to design and implementation.

Planning - Definition

Because there are many reasons for developing a Web site, the first step in the planning process is to define clearly the primary objectives, target audiences and project scope. Although Web sites may be developed to create a presence on the Web, this reason alone is not sufficient basis for an effective site. Effective Web sites work because they achieve substantial and well-defined objectives. Taking time to clarify a site's objectives, to analyze its audiences and to develop strategies for appealing to each audience is necessary for the success of the whole project.

To begin, a one- or two-sentence Web site mission statement should be written that summarizes the goals of the site and its desired effect on the audiences served. With a general mission defined, the next step is to develop specific strategies for achieving these goals. This may be as simple as a list of the three things that the site should accomplish, or it may involve more detail that outlines numerous objectives. Either way, it is important to keep the objectives specific and realistic. Determining the measurement by which achieving objectives will be assessed also is crucial. Will it be determined by the number of visitors? By which sections are accessed most? By critical acclaim? The answer is best decided by each project.

In the definition phase of NCPTT's Web site redesign, five objectives were formulated. The NCPTT Web site is intended to — 1) fulfill goals and objectives outlined in NCPTT's mission and long-range strategic plan; 2) serve as a clearinghouse and delivery mechanism for information sponsored, collected and developed by NCPTT; 3) establish and promote NCPTT's role in the conservation and preservation community; 4) promote online communication for individuals and organizations involved in preservation and conservation; and 5) provide a means to measure the impact and reach of NCPTT's work.

Another vital issue to be addressed in the definition phase is a Web site's audiences. Web sites typically address five potential audiences — two public, two private, and one semi-private (Figure 1). A site's public audiences include a target audience and the general browsing public. Private audiences include internal users of the site who may have access to proprietary sections — such as an organization's staff — and administrative users who are responsible for maintaining the site and its content. Finally, a site may also have a semi-private section for special users, such as an area available only to registered users. The distribution of audience sizes will vary among projects. Some sites may have a very wide target audience and a fully developed internal Web site for staff; other sites may have a very narrow target public and an administrative section used by only a few people.

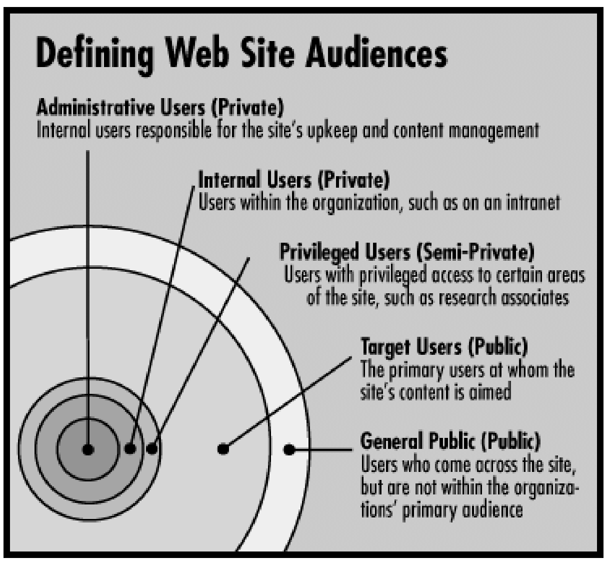


Figure 1. Potential Web site audiences.

Where possible, it is important to determine as much as possible about each audience, such as reasons for visiting the site, anticipated frequency and length of visits, the educational and professional background of the audience and whether the user will be accessing the site from home or office. While it is acknowledged that ascertaining this kind of information is difficult, it can be valuable to the design process if it is available. Other useful demographics are related more to the audience's hardware and software configurations, such as computer platforms, browser software and speed of Internet connection. If the project involves redesigning an existing Web site, server logs that record information about Web site visitors may provide some of these statistics.

Since NCPTT's project involved redesigning a Web site that functioned primarily as an interface to NCPTT's gopher — and since there were no server logs available — we were not able to generate target audience statistics. However, NCPTT's audiences were easily defined. The target audience is professionals in the fields of archeology, historic architecture, historic landscapes, objects and materials conservation and interpretation. NCPTT staff, a subset of the broader preservation community and the general browsing public were also considered in the planning process.

Planning - Architecture

The second phase of site design — architecture — pairs the objectives and audiences with content. Methods involved in choosing appropriate content, organizing and prioritizing content and creating a clear navigation system through the content are drawn from the growing field of information architecture.[3] When designing the information architecture of a site, use the already defined objectives and audiences to decide the nature of the site's content. Actual content might come from sources such as a digital version of a museum's collection or a printed history of a preservation organization. Other content may be developed specifically for the site, such as a database of an organization's members or an online discussion forum.

For NCPTT's Web site redesign, audience needs were determined to fall into three categories: information about NCPTT for those who may be unfamiliar with the organization, information about programs sponsored by NCPTT and current and archived information on preservation topics. Actual content was drawn from several sources. Information about NCPTT and its programs was adapted from existing materials. A new online version of NCPTT's newsletter, *NCPTT Notes*, was adapted from print versions of the publication. Additionally, NCPTT already provided a wide variety of preservation-related information via its existing gopher site or via hardcopy, such as research findings, conferences, job postings, funding opportunities and other online resources. This information continues to be relevant and has been transferred to a searchable database system for enhanced access via NCPTT's site.

Content organization is the important next step. While there are few universal rules for information design, taking the time to examine three principal factors leads to better results. First, determine natural organizational systems within the content that would help to make the information more accessible. For example, the information presented in a Web site on southwestern archeological projects might be organized geographically or chronologically. Second, determine particular objectives that would require giving certain information priority. For NCPTT's Web site, disseminating results of PTTGrants and

PTTProjects is critical, so that information will be prominently featured and available via various routes. Third, determine what the site's audience will be looking for when contacting the site.

Working from the three categories of audience needs, NCPTT's Web site is configured in three "meta" components — "About NCPTT," with generally static organizational and background information; "NCPTT Components," a more active section for NCPTT's three components — research, training and information management; and "Resources," a searchable database system of preservation-related information that is currently in development.

Visualizing the information architecture of the site is best done with a site schematic — a graphic representation of the site's structure that shows how the information in the site is organized and connected. Figure 2 is a corner of the very large NCPTT Web site schematic; it is impossible to reproduce it in its entirety without its being indecipherable. The site schematic also is a good tool for planning future areas of the site, since these can be placed into the master scheme even if they are not developed initially. NCPTT's site schematic encompasses all planned functions of the site. During the phased development process the plan is fine-tuned to deliver substantive information to an appropriate audience.

Development

Upon completion of the planning phases, the three phases of development — design, implementation and integration — can begin. However, elements of the planning process will continue throughout the development phases. We recommend that development takes place in a staged approach to facilitate modifications of the plan and so that each stage can build upon the previous.

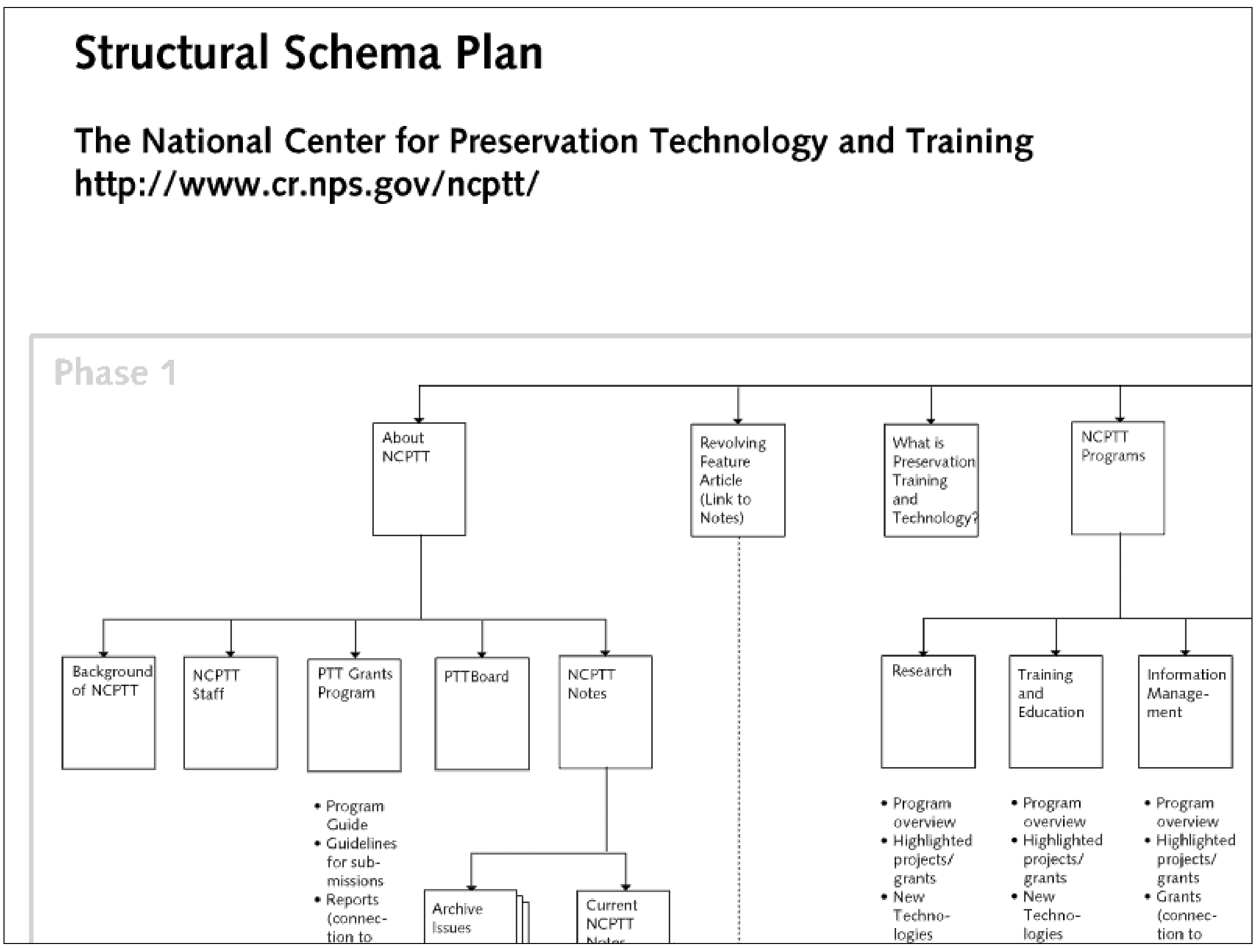


Figure 2. NCPTT Web site schema.

Development of NCPTT's Web site was divided into four phases. Phase one encompassed the design of the graphical user interface and development of static site content, including descriptions of NCPTT, its mission, its program areas and the Preservation Technology and Training Board.

The design of the graphical user interface began by laying out the site's organization and testing the information architecture with basic navigation elements. NCPTT staff reviewed the navigation process for ease of use. Next, a few elements were put in place (Figure 3) and then sample images and text were added giving it an appearance close to the final product (Figure 4). Figure 5 graphically represents the main concept sections of the core page. The final core page can be viewed at www.ncptt.nps.gov; the center image varies according to the current feature.

The structure of second level pages — NCPTT Programs, About NCPTT, Resources — maintains the same navigational pathways and graphical interface as the core page. See Figure 6 for a graphical representation of the second level page sections. A final second level page can be viewed at www.ncptt.nps.gov/im. (Figure 7) All pages at this level will have the same look and navigational tools; only the details in text and images will vary.

Phase two included designing the databases that will replace the gopher-based Resources section and implementing one prototype database — Preservation Internet Resources, which replaced the annotated list known as "Internet Resources for Heritage Conservation, Historic Preservation and Archeology". A thorough and exhaustive needs analysis was undertaken by the Web site designers. NCPTT staff were asked to complete questionnaires for each category of information to be disseminated via the Web databases. Staff were also interviewed in person by the development team. The goal was to address the current scope of the database project and to insure the scalability of the system developed. Design specifications for all of the tables in the system were codified in a 78-page document entitled "Design Documents for the NCPTT Web Site: Phase II Development," which provided the basis for Phase three.

Phase three involves implementing the other databases in the system. The system includes General Databases (conferences, jobs, grants); Program-related Databases (Training and Education, Analytical and Materials Testing Directory, Materials Research Bibliography); and Grants, Projects and Publications Databases (Grants and Projects Catalog, Publications Management).

Phase 4 is intended to add functionality to the Web site and to target a smaller audience — a PTTCommunity. Not yet through its planning stage,

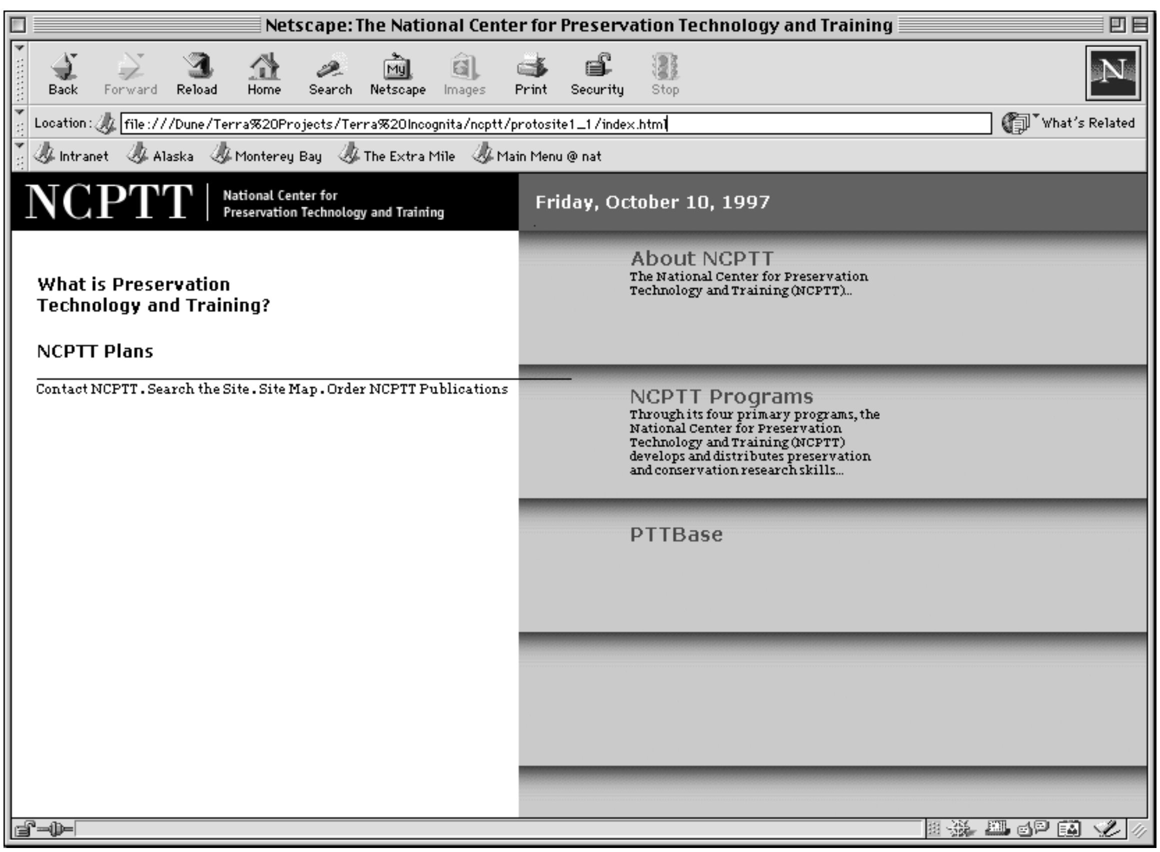


Figure 3. Basic elements of NCPTT Web site.

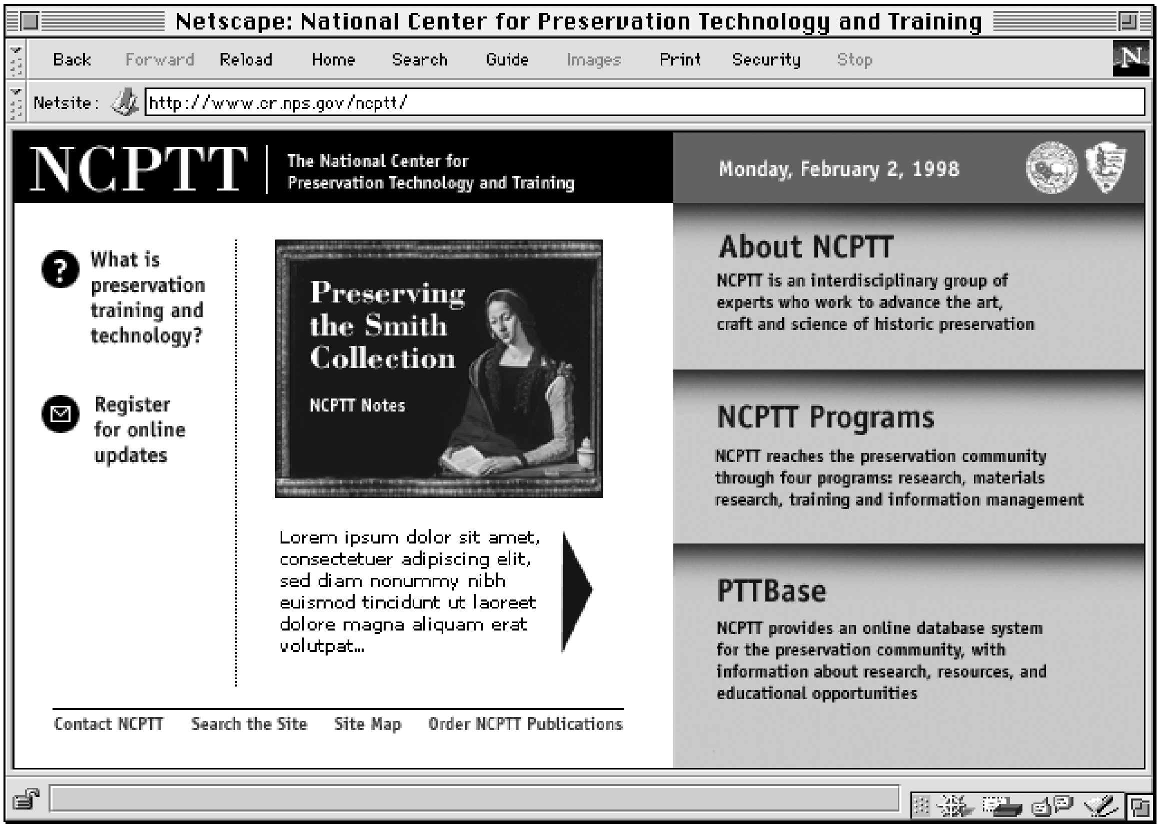


Figure 4. Sample images and text.

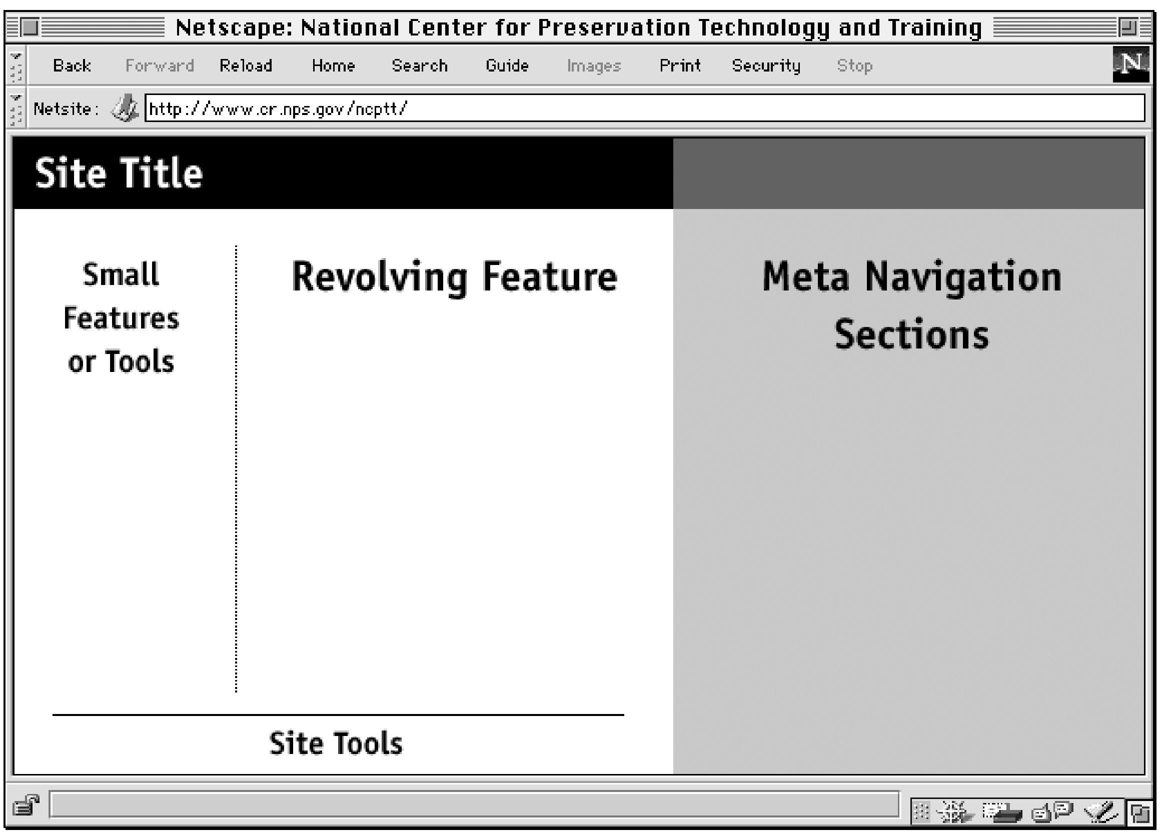


Figure 5. Main concept sections.

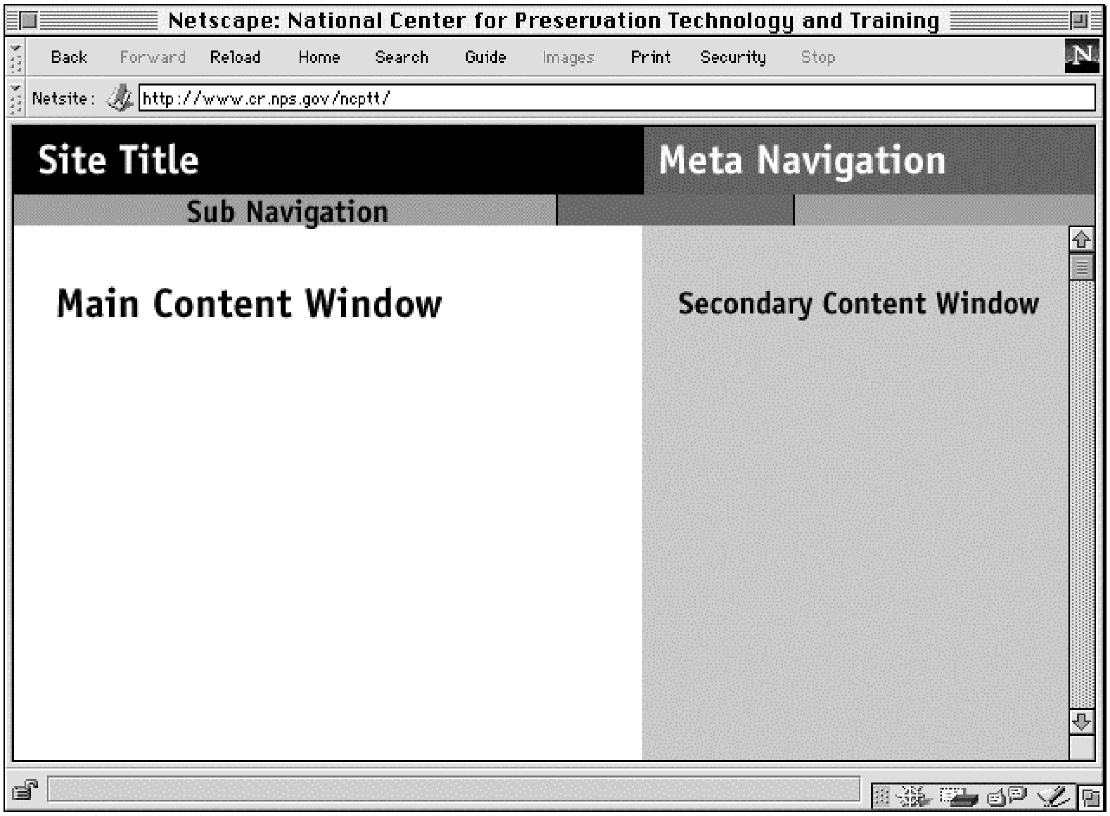


Figure 6. Graphical representation of second level sections.

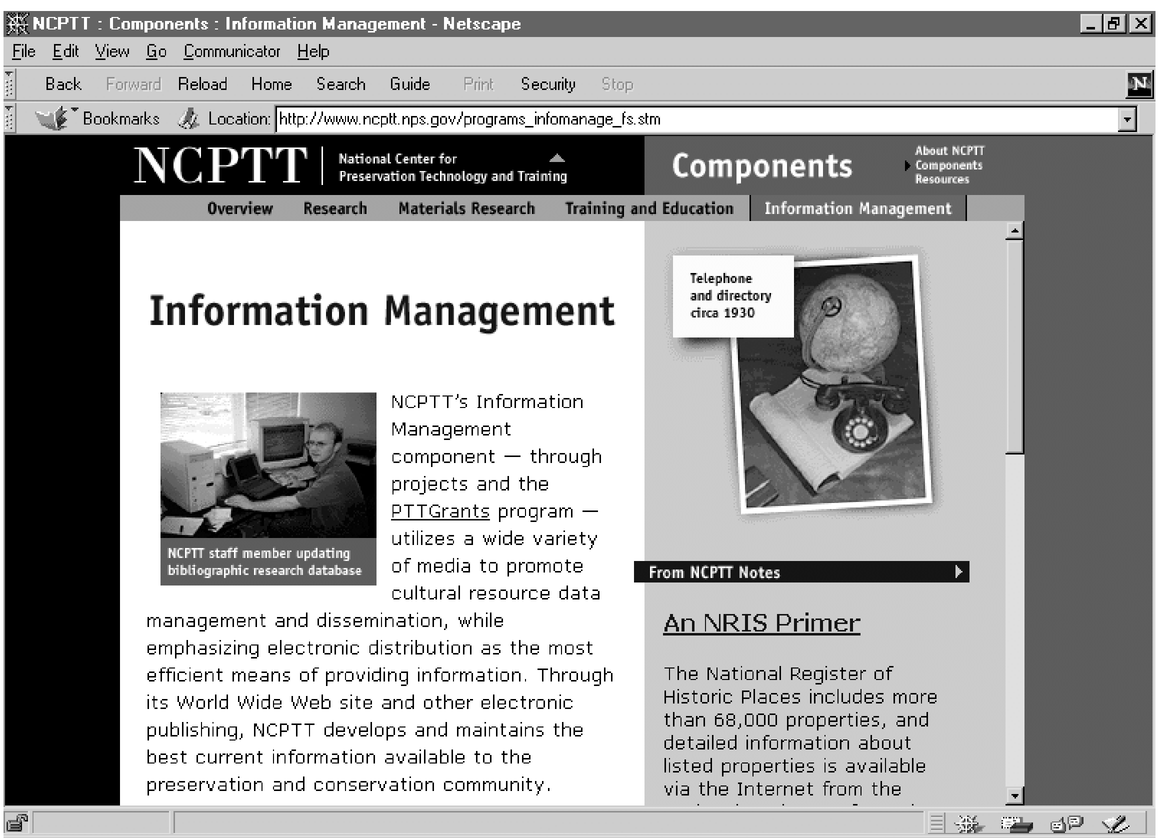


Figure 7. Final second level page – Information Management component.

Phase 4 may include a system that will allow users to register to be notified when items matching their interests are added to NCPTT's Web site and may also allow users to input additions to the database system for consideration by NCPTT staff. In addition, a system to allow online submission of PTTGrants proposals and to allow PTTGrants reviewers to access proposals and submit comments online will be developed.

Conclusion

The planning phases are the real foundation upon which a successful Web site is built. The surface of an archeological site usually gives some indication of what lies below — what could be called the "content" of the site. As we all know, it's necessary to remove the soil to discover the "content." But the surface of an archeological site will rarely, if ever, indicate everything that is below. There are often unexpected discoveries — "content" that was not anticipated based on the surface survey or collection. And since archeological sites are created over time through natural processes, obviously no one plans or develops surfaces to show the underlying "content."

However, we can control what could be called the Web site's "surface" — the core page of the site. And unlike an archeological site, the full content of a Web site can — and should — be clearly mapped at the "surface." There should be no surprises to users, especially unanticipated lack of content.

Notes

1. Mary S. Carroll, "Internet Offers Many Electronic Journals of Preservation Interest," *NCPTT Notes* 15 (1996): 8-9.

Mary S. Carroll, "The National Archeological Database," *NCPTT Notes* 17 (1997): 6-7.

Mary S. Carroll, "Preservation Internet Discussion Forums," *NCPTT Notes* 21 (1997): 6. <www.ncptt.nps.gov/notes/21/5_fs.stm>.

David Carlson, this volume.

S. Terry Childs, this volume.

John Hoopes, "Avoiding the Driest Dust that Blows: Web Site Reports," *SAA Bulletin.* 17.1 (1999): 23, 26-27, 39.

John Hoopes, "You've Got News! Archaeology Journalism on the Internet," *SAA Bulletin.* 17/2 (1999): 26-28.

John Hoopes, "Electronic Quipus for the 21st Century: Andean Archaeology Online," *SAA Bulletin.* 16.1 (1998): 20-22. <www.anth.ucsb.edu/SAABulletin/16.1/SAA16.html>.

John Hoopes, "The Online Lab Manual: Reference Collections on the Web," *SAA Bulletin.* 16.5 (1998): 17-19, 39. <www.anth.ucsb.edu/SAABulletin/16.5/SAA14.html>.

2. Bart Marable, "Once Upon a Time: Using New Narratives in Educational Web Sites" (Paper presented at Museums and the Web 1999, New Orleans) <www.archimuse.com/mw99/papers/marable/marable.html>.

Bart Marable, "Bringing Stories to Life Online," *Web Techniques*. 4.3 (1999): 18-23.

Bart Marable and Raymond Simmons. "Accurate Wayfinding With Javascript," *WEB Techniques*. 2.7 (1997). <www.webtechniques.com/archives/1997/07/junk/>.

3. Clement Mok, *Designing Business: Multiple Media, Multiple Disciplines* (San Jose: Adobe Press, 1996).

Louis Rosenfeld and Peter Morville, *Information Architecture for the World Wide Web* (Cambridge: O'Reilly and Associates, Inc., 1998).

Federal Archeology on the Internet: Current Status and Future Directions

S. TERRY CHILDS

The Internet provides a powerful and dynamic tool to inform a wide range of audiences about archeological projects, programs and interpretative results; provide scholars and students with interactive access to research databases and other materials; and provide cultural resources management staff with various management tools. Federal agencies increasingly use the Internet for some or all of these functions, but with considerable variability due to a number of factors. This paper examines current efforts to bring Federal archeological activities and resources to light on the Internet and explores future directions and possibilities.

I want to begin by clarifying what I mean by "Federal archeology". It is the management, preservation and protection of archeological resources on and from Federal lands. This includes inventorying sites; conducting CRM-based and research-based archeological projects; preserving and protecting known sites in place; dealing with the curation of the collections, records and reports resulting from archeological projects; writing relevant laws and regulations; helping to implement those laws; and providing public education and outreach about archeology on Federal lands. It is important to understand that archeologists work for many different government agencies, each with different missions that relate to both archeology and education. Federal archeologists also work with many others at the state, tribal and local levels, as well as in academia.

As such, Federal archeology and related cultural resource programs definitely have made use of the Internet, although with significant variation across agencies. I begin by summarizing the primary uses and then point out some problems related to them. The second part of this paper examines some future directions for effective use of the Internet for Federal archeology, especially if some of the current bumps on the cyber-road can be overcome.

Federal Archeology on the Internet

The two primary parts of the Internet used by Federal archeologists are listservers and the World Wide Web. Listservers such as Arch-L, Histarch-L, AIA-L, Anthro-L and Museum-L have become important vehicles of communication that have encouraged Federal archeologists, often isolated in parks or forests, to participate in both theoretical and practical discussions of current issues, to circulate new initiatives and to ask for assistance from unknown colleagues. Some groups of Federal archeologists and related colleagues have initiated their own listservers, such as the one for the Federal Preservation Forum. These serve to promote communication about in-house cultural resource issues that cross-cut Federal agencies and disciplines, including historic preservation, cultural landscapes, American Indian consultation, deaccessioning archeological collections, and Sections 106 and 110 compliance.

The other major use of the Internet over the last several years is the Web. Today the Web serves a variety of functions for a variety of audiences. The most common materials provided by Federal agencies are for basic public outreach. These typically include descriptive information about a particular Federal archeology program or activity, such as the cultural resources programs of the Bureau of Reclamation[1], the Fish and Wildlife Service[2], or about Federal archeology in general.[3] Sometimes a division or group in an agency uses the Web to advertise their expertise and technical capabilities, perhaps in hopes of some future partnership activities, such as the Army Corps of Engineers' Construction Engineering Research Laboratory.[4] Much of this type of information is reprocessed print material that focuses on "who we are" and "what we do".

The Web also is used for access to existing documents about Federal archeological initiatives, cooperative agreements, guidelines, planning, law enforcement, and training. The US Army, for example, has a useful one-stop shopping compilation of its cultural resources documents[5] that were once only available in print. Other materials that once took real time to order and receive, but are now readily available on the Web, are grant information and applications. Some good examples include the National Science Foundation,[6] National Endowment for the Humanities,[7] and the Native American Graves Protection and Repatriation Act grants for museums and tribes.[8]

Federal archeologists and other cultural resources professionals often feel compelled to provide access to the many statutory requirements that justify and structure their work. The most extensive list is provided by the National Park Service.[9] A real plus of the Web, however, is the ability to offer ready access to potentially significant draft legislation that used to be difficult to find. This service has been provided on the Web by the Advisory Council for Historic Preservation and the National Park Service on various occasions. Additionally, the Advisory Council recently began an interactive online forum on Federal historic preservation to gather comments on related Federal legislation and its effectiveness[10] (Figure 1).

Federal staff also are taking advantage of the Web to distribute widely used, free magazines and publications, such as *CRM*[11] (Figure 2) and the technical bulletins of the National Register of Historic Places,[12] to a broader audience than has ever been possible. Not only are current issues of *CRM* posted on the Web as they come out in print and sent to a limited distribution, but twenty years of back issues now are available, indexed and searchable in a database.

Public outreach on the Web involves enhancing education about Federal archeology for students beginning as early as kindergarten and through college and beyond. This usually consists of pages on the excavations and interpretation of a site such as the Five Points site in New York City by the General Services Administration (GSA)[13] or African-American households at Manassas National Battlefield (NPS).[14] Whereas some of this type of material may be acquired from brochures and other printed material, the Web offers a relatively cheap way to provide additional color photos of the context, excavation and artifacts resulting from a project that is not possible in print.

Federal archeologists, curators and other CRM specialists are using the Web to provide information on educational and volunteer services, such as at the Bureau of Land Management's Anasazi Heritage Center[15] and the US Forest Service's Passport in Time in particular forests (i.e., Tahoe National Forest).[16] Most states now designate one week or month each year to promote local and statewide understanding and participation in archeology and many have active Web sites on these programs. The National Park Service provides an invaluable compilation of all the states with archeology weeks or months, along with contact information and active links to the state Web sites. [17]

While the Web is used to increase awareness of such programs, other Federal archeologists have begun to develop educational products targeted to particular audiences on the Web. These include "Ancient Architects of the Mississippi",[18] the "Teaching with Historic Places" lesson plans — some having to do with archeological sites[19] — and virtual tours of archeological sites in a region or by state.[20] Another type of education tool is a virtual museum exhibit around a particular theme. The exhibit on Civil War camp life at Gettysburg,[21] for example, focuses on the daily lives of young men at war with poignant similarities between the objects of yesteryear and today (Figure 3).

Federal archeologists and cultural resource specialists also use the Web to provide unique services, such as interactive databases that allow researchers to search for and explore particular interests and topics. The Reports module of the National Archeological Database,[22] for example, provides bibliographic information on archeological projects conducted across the United States, particularly from the gray or unpublished literature. Another important NPS database is the Native American Consultation Database[23] that provides tribal contacts for a variety of issues, such as NAGPRA and unanticipated discoveries. The US Information Agency's International Cultural Property Protection program[24] offers an image database of pillaged artifacts subject to import restrictions (Figure 4). As well, considerable value can be added to basic information found in other venues and media when organized and standardized in a searchable database, such as the Preservation Internet Resources database[25] of the National Park Service's National Center for Preservation Technology and Training. Some of these databases were first made accessible on the Internet via telnet or gopher, but have been converted to the more user-friendly formats permitted on the Web.

Geographic Information System maps are another means to summarize, analyze and present complex information in graphic form on the Web, usually for professionals. The Multiple Attribute Presentation System module of the National Archeological Database,[26] for example, offers a useful set of national level GIS maps. Of particular interest is the capability to create a unique data layer on a specific archeological phenomenon in order to study its regional or national distribution, such as the national distribution of Paleoindian projectile points.[27] The results of extensive local research using GIS can be communicated effectively by taking advantage of the color graphics capabilities of the Web, such as identifying and assessing the earthworks at battlefield parks.[28]

Problems

While Internet-based materials have considerably expanded the public's exposure and access to the work and products of Federal archeologists, there are a few problems and issues that must be acknowledged before we look to the future of Federal archeology on the Internet.

The first is audience. It is difficult to design and create materials for both the general public who pays their taxes and are interested in archeology yet often access the Web via modem, and professional colleagues who want very different materials and often have direct connections to the Web. A second factor is frequency of visits — the casual browser who unexpectedly finds interesting or useful information on a Federal archeology site versus the repeat user. As a

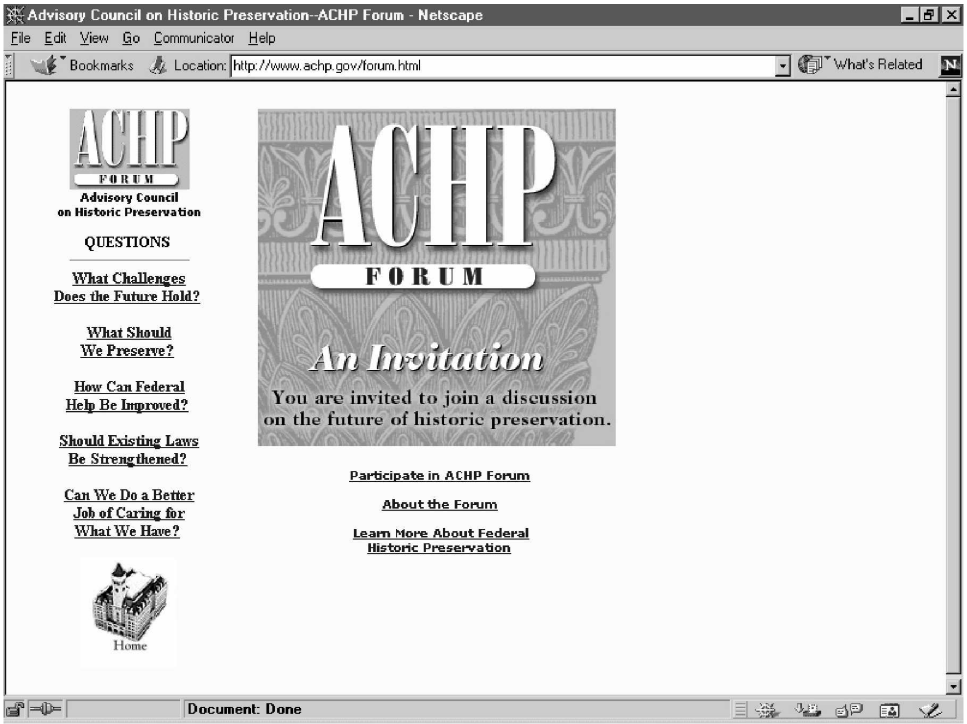


Figure 1. Advisory Council for Historic Preservation forum on historic preservation.

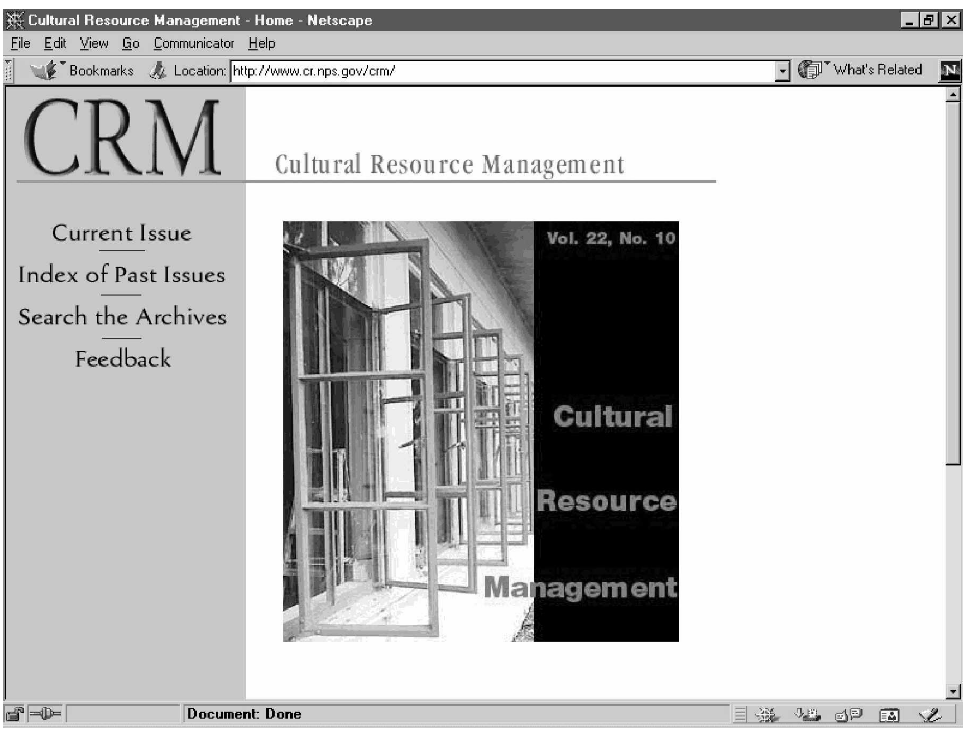


Figure 2. *CRM* magazine Web site.

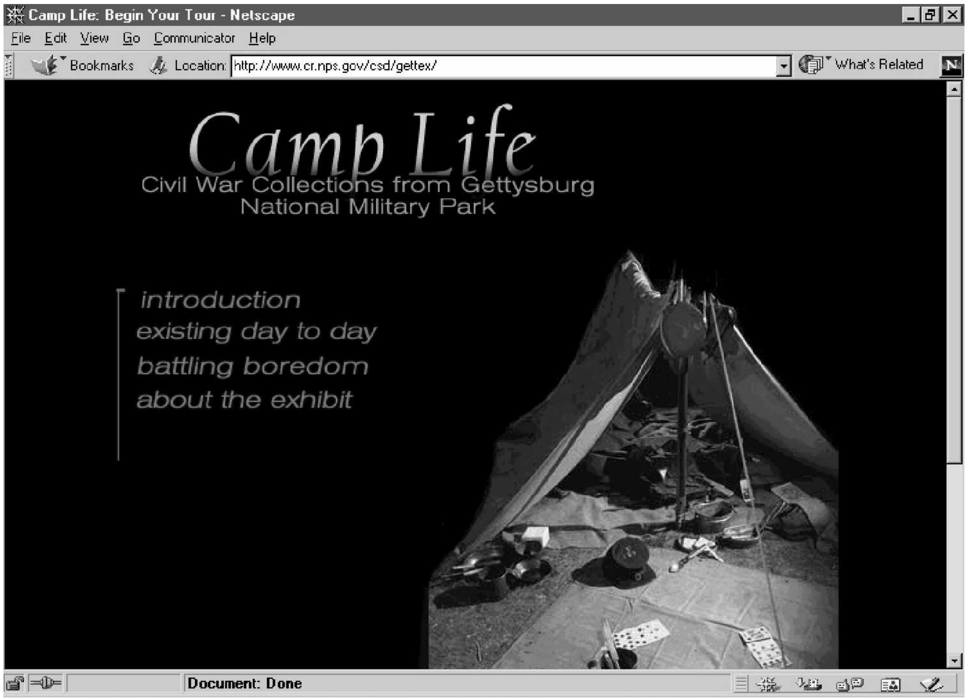


Figure 3. Civil War camp life Web exhibit.

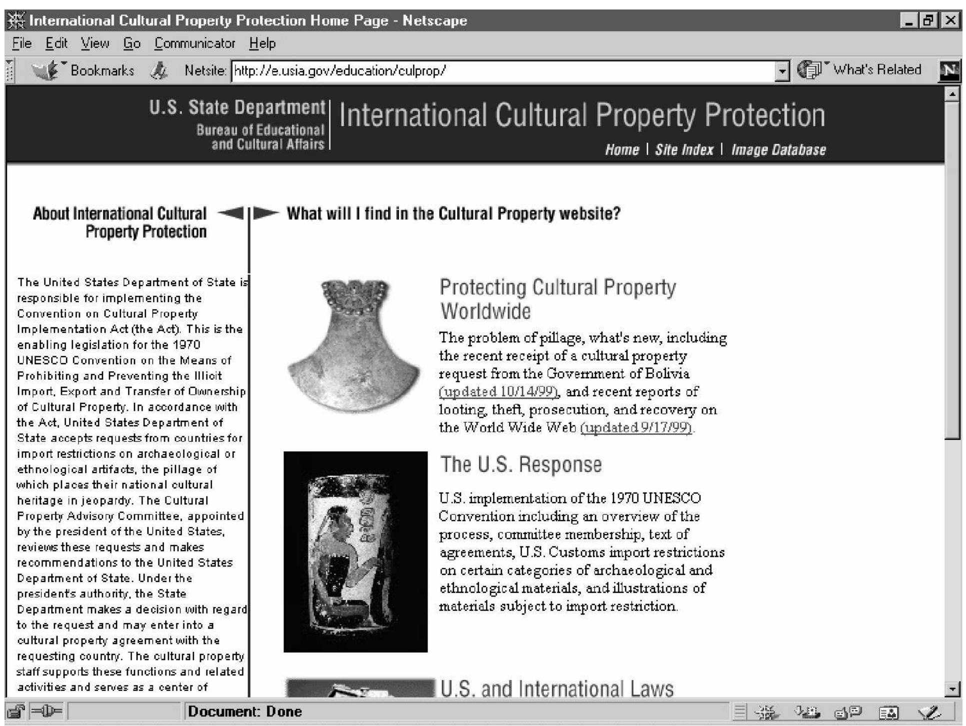


Figure 4. International Cultural Property Protection program - U.S. Information Agency.

consequence, many Federal Web sites on archeology are a hodgepodge of materials for different audiences. The users, in the end, really need to know what they want at a Federal Web site and where to go before they begin.

Finally, a related problem is accessibility to desired information. It is sometimes difficult to find the materials on archeology at many Federal Web sites because the overall site lacks a good organizational structure that recognizes archeology or even cultural resources. A few Federal Web sites lack a search engine or index. Furthermore, individual Web pages may lack good metadata (documentation about the content) to enhance quick retrieval by search engines.

Long-term Web site development and maintenance is another issue. Many Federal archeology Web pages have dated content or bad links. The bottom line is that most Federal agencies with archeology programs do not have a formal Web design and development infrastructure for cultural resources, let alone archeology. What is on the Web, therefore, often is the product of one or two interested and hard-working archeologists who do the best they can when they have time. Regular maintenance is difficult and new product development without extra help and money is even more difficult. It is easier to reuse relatively static, printed materials for wider distribution on the Web rather than create new materials for this challenging medium. For interactive databases, there is the serious issue of regular updates and maintenance.

Finally, Federal managers have been relatively slow to understand the powerful, far-reaching benefits of the Web for communication and education about their resources, products and programs. Therefore, they have not invested in creating a long-term infrastructure to develop and maintain their Web sites since this involves both money and staff that are scarce valuable commodities. This is slowly changing and with it is some increased understanding of basic Federal responsibilities to their users. These responsibilities include access for handicapped users, copyright, photo release permissions and standardized metadata to document the sources and content of Web materials for long-term use. The latter will become particularly critical as collections of materials, such as archeological reports, archival documents and archeological collections are digitized and GIS data layers are created and posted on the Web.

Where Federal Archeology Might Go on the Internet

The National Park Service's Cultural Resources Web site, "Links to the Past,"[29] now handles about 9,000 users per day at an average of 15 minutes per session (Figure 5). Given these numbers, NPS managers are beginning to invest some money and staff expertise in the Web to utilize its unique capabilities and begin to take it in new directions.

If Federal managers do invest further in the medium, there are a number of new ways that Federal archeology can take advantage of the Internet in general. However, given limited resources and constant changes to information technology, it is becoming imperative to work in partnership with other Federal agencies or non-Federal organizations on a particular project to take advantage of their in-house expertise. This is already happening. For example, the National Park Service's Archeology & Ethnography Program works with the Center for Advanced Spatial Technologies at the University of Arkansas on database and GIS development and the Bureau of Reclamation is working with the Archeological Research Institute at Arizona State University as Peter McCartney discusses in this volume.

My discussion on new directions revolves around a key factor mentioned above — audience. Which audience warrants investment — the interested public or fellow professionals and their students? the frequent user or the casual browser who discovers a useful site? I suspect that some agencies will pursue the public outreach and education route, others will pursue the professional route and a few will attempt to cater to all.

General Public

For the general public effective interpretation must involve highlighting the history and prehistory of the sites preserved and protected on Federal lands and showing the relevance of those sites to today's life. One direction may be to construct meaningful connections to the past through links to Web sites that deal with modern life issues, businesses and recreation. For example, Federal archeologists who study ancient and historic period ironworking and create Web pages about that research might contact the Web managers of modern steel company sites. They can argue that it is good for public relations to set the development of steel making in its historical context and perhaps work to develop a joint project. Similar arguments can be made about the connections between large-scale corn farming and the rise of corn-based agriculture in the Americas or the connections between major weather storms, lighthouses and shipwrecks.

Another direction relates to the economic value of archeology, which is of great interest to community leaders, tourist bureaus, newspaper reporters and television producers. In particular, the growing heritage tourism industry can be connected intimately to archeology through the prehistoric sites, historic buildings and museums to be visited. The Web provides ways to expose the public to such relationships, to set the historical context of the sites to

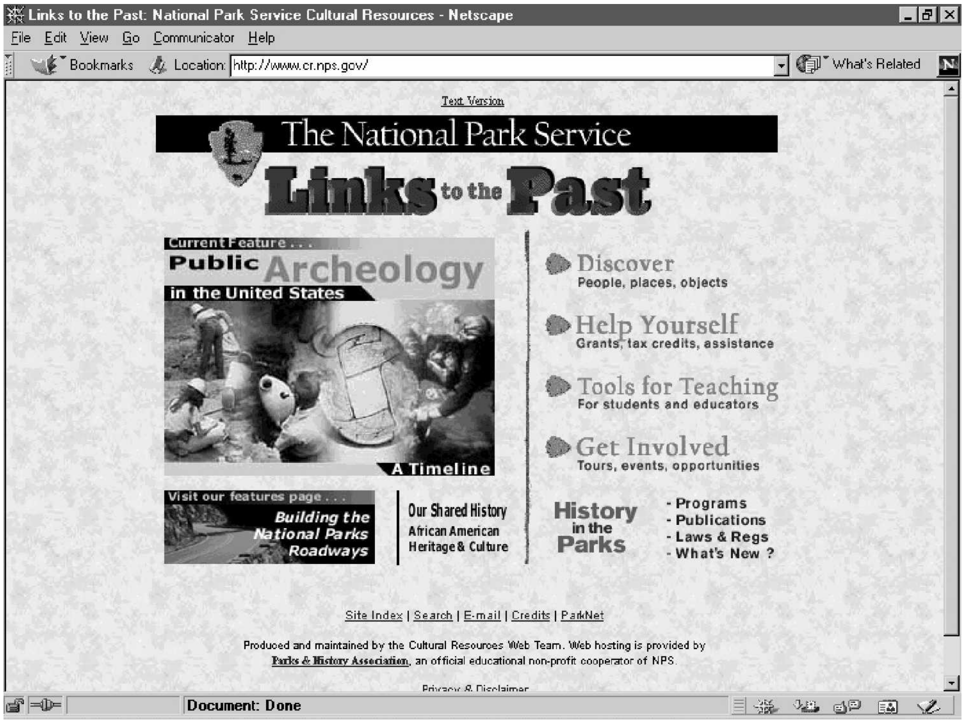


Figure 5. Cultural resources area of the National Park Service's Web site.

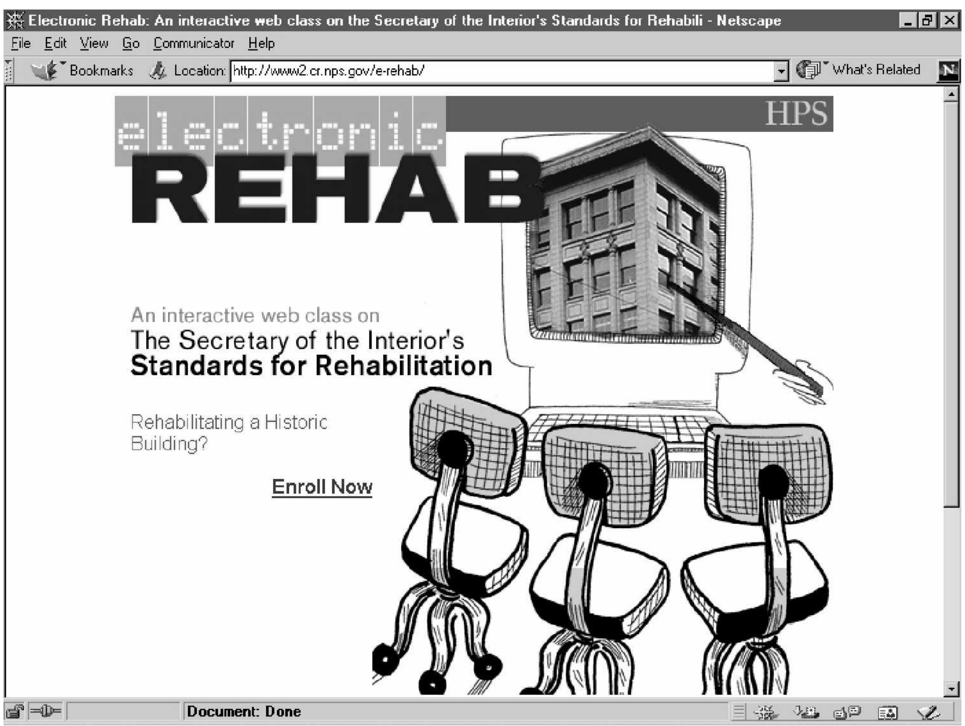


Figure 6. Electronic Rehab -- National Park Service educational module.

be visited and to develop significant partnerships with state and local groups through thematic travel itineraries and features. Dynamic maps allow a Web user to visualize a region and its road networks, click on a series of recommended sites and receive useful information (written and/or audio), photographs and videos about them. The National Register of Historic Places has constructed a series of itineraries on cities, like Chicago and Seattle, and on themes, such as the Underground Railroad and women's history. Similar efforts may begin to focus on archeology themes, such as early efforts on "Ancient Architects of the Mississippi"[30] and the archeology in the "Golden Crescent" of the southeast.[31]

Another focus is to highlight the process of discovery during an archeological excavation for the general public. The Web now offers the possibility to document an excavation in almost real time using on-site digital cameras, as well as written descriptions of daily progress. Although this sort of educational endeavor costs substantial money for equipment and staff expertise, it has proven a successful and lively way to engage the public in archeology. Not only does discovery excite the public due to the unknown possibilities, but a daily account of an excavation can teach about archeological techniques, decision making and interpretation. A first attempt has been made by the Southeast Archeological Center of the National Park Service at Cumberland Island National Seashore.[32] Federal archeologists do develop research endeavors, such as the National Historic Landmark Underground Railroad Archeological Initiative. Perhaps Federal archeologists will budget for or seek outside funding for such Web efforts during early project planning.

Professionals

There are a number of new directions that Federal archeology can follow for its professional audience, including academic, CRM and amateur archeologists, as well as other CRM specialists. These can be lumped into three main categories: professional education, CRM responsibilities and research.

In terms of professional education, a primary new direction for Web-based Federal archeology may involve distance learning modules. This is not just putting up a syllabus, some links to Web sites related to the subject matter and a bibliography. It is developing short training modules or a full course on a topic that is not standard to graduate archeology programs, yet is fundamental to the work of Federal archeologists. These might include the legislative history of Federal archeology, archeological ethics, archeological curation, object conservation and project management. An excellent example of such an endeavor that is used by hundreds of individuals per week is the NPS educational module for preservationists called "Electronic Rehab"[33] (Figure 6). Note that these online courses will probably never develop into interactive, credited courses unless developed in partnership with a university that can monitor and evaluate performance and give course credit. The modular format, however, can be readily incorporated into university teaching.

Another category of future Web development relates to providing a national perspective on basic CRM responsibilities occurring at the Federal, state, tribal and local levels. Many Federal, state and tribal archeologists create, work with or have access to large amounts of data generated through compliance with Federal laws and regulations. These include archeological site records, nominations to the National Register of Historic Places, archeological project reports, inadvertent discoveries related to NAGPRA and Federal collections and associated documents. I expect that most new efforts to deal with and provide access to information will be in the form of searchable databases including digitized images, GIS mapping and combinations of the two.

Not only do these data cover activities from large tracts of land, but they often come from multiple sources. While databases provide a mechanism to gather large quantities of standardized data for a variety of uses and the Internet facilitates access to that data, the data contents often need to be validated and kept up-to-date. I believe a new direction for Federal archeology and CRM work is developing online data entry capabilities to handle these needs. For example, the Reports module of the National Archeological Database is now several years out of date since the National Park Service infrastructure to handle data updates crumbled upon staff reorganization. In response, an online data entry system is being explored to allow CRM contractors to enter the citation information about their reports into the system. Technically, this is relatively easy to do these days. The hard part is staffing for data entry and validation. Therefore, a validation system is being developed where the State Historic Preservation Offices, who currently supply bibliographic data to NADB-Reports, can be periodically notified about newly created records. Designated SHPO staff would review the records with full editing capabilities, submit each record to the master database for immediate upload and then have the option to download that same record to their local database. Given that this process involves SHPO offices in over 50 states and territories with their own staffing and priorities, there are numerous organizational hoops to overcome. Careful analysis of workload issues and hardware and software compatibility must accompany efforts to implement an online data entry system at the national or even regional levels.

Another type of potential information system for Federal archeology on the Web involves using GIS to access and analyze specialized information in map formats. For example, there is increasing need to assist CRM professionals, as well as local law enforcement, in locating appropriate tribal contacts for NAGPRA related issues or inadvertent discoveries of historic and prehistoric burials. A GIS interface to the Native American Consultation Database[34] using data layers such as state, county, rivers and tribal reservation boundaries, at a minimum, would facilitate the discovery of appropriate tribal contacts in times of immediate need. Because GIS can provide precise locational information, Federal archeologists must be careful to consider the requirements of laws, such as the Archeological Resources Protection Act and the Electronic Freedom of Information Act, when preparing GIS maps for online use.

Also related to GIS on the Internet is developing the ability to build maps online from a collection of public domain data layers related to archeological site density, the environment, historic census records and other relevant information. Armed with choices from a large number of data layers, perhaps broken down by state or ecosystem, users could facilitate decision-making on archeological resource management, predictive models and research. This capability is being developed at the state level in Arkansas[35] and could be applied on a national basis, including at particular national parks or national forests.

The final category related to possible future developments on the Internet deals with research, as well as CRM responsibilities. Here we focus on a source of archeological data that is often underused, primarily due to accessibility problems — collections, records and reports. We have discussed the attempt to make "gray literature" reports more widely known through the Internet, but what about the huge number of archeological artifacts and documents owned by the Federal government — well over one billion? There is one effort by the National Park Service to put online summary information of museum collections housed in all their national parks and regional centers.[36] This information, however, is not detailed enough to determine the nature of particular collections in a repository in order to facilitate the development of a research project.

Two sources of pressure to make collections more accessible and accountable may stimulate Federal repositories and the non-Federal repositories that care for Federal collections to develop online searchable databases of their collections. One source is the White House and other government groups, who initiated Save America's Treasures[37] to provide needed care for significant Federally-owned collections, among other cultural resources, and to educate the American public about these hidden treasures. The other source is professional archeologists who are interested in access to collections information for research purposes. Search fields in repository databases might include the source project name, source project location, the range of materials in a collection, cultural affiliation and condition to facilitate research project planning for professionals.

Another aspect of archeological collection research that might receive some attention in the future is the use of three-dimensional imaging and photogrammetry. Although it is unlikely that Federal archeologists will be heavily involved in the further development of this technology in general, they can certainly benefit from supporting the development of its use on the Web. If researchers can examine a whole object in three dimensions from home or office, do basic measurements and examine decorative style and basic technological features, then they do not necessarily have to visit the museum in which the objects are housed. This frees repository staff to pursue activities other than supervising researchers.

Conclusions

There is a considerable amount of material related to Federal archeology now on the Internet, but it is primarily descriptive and is not widely known. It also does not help that this information often is difficult to find and poorly planned and organized. But Federal managers are beginning to better understand the strengths of the Internet for communication, education, the relative ease of data sharing and accountability to a vast public and are improving Federal Web sites. It may take some time before many of the future directions I described are fully invested in and implemented, but some new footsteps are being taken on the Internet cyber-road.

Notes

1. <www.usbr.gov/cultural/>

2. <refuges.fws.gov/NWRSFiles/CulturalResources/CulturalResources.html>

3. <www.cr.nps.gov/aad/fedarch.htm>

4. <www.cecer.army.mil/>

5. <aec-www.apgea.army.mil:8080>

6. <www.nsf.gov>

7. <www.neh.fed.us/html/applying.html>

8. <www.cr.nps.gov/aad/nagpra.htm>

9. <www.cr.nps.gov/linklaws.htm>

10. <www.achp.gov/forum.html>

11. <www.cr.nps.gov/crm/>

12. <www.cr.nps.gov/nr/nrpubs.html>

13. <r2.gsa.gov/fivept/fphome.htm>

14. <www.nps.gov/rap/exhibit/mana/text/rhouse00.htm>

15. <www.co.blm.gov/ahc/teach.htm>

16. <sv0505.r5.fs.fed.us:80/tahoe/tnf_vol_pit.html>

17. <www.cr.nps.gov/aad/statearc.htm>

18. <www.cr.nps.gov/aad/feature>

19. <www.cr.nps.gov/nr/twhp>

20. <www.mwac.nps.gov/where_to_see/>

21. <www.cr.nps.gov/csd/gettex/>

22. <www.cr.nps.gov/aad/nadb.htm>

23. <www.cr.nps.gov/aad/nacd/>

24. <e.usia.gov/education/culprop/>

25. <www.ncptt.nps.gov/pir>

26. <www.cr.nps.gov/aad/nadb.htm>

27. <www.cast.uark.edu/other/nps/maplib/USfluteddens.html>

28. <www2.cr.nps.gov/gis/reports/fishhook/intro.htm>

29. <www.cr.nps.gov>

30. <www.cr.nps.gov/aad/feature/>

31. <www.cr.nps.gov/goldcres/>

32. <www.cr.nps.gov/seac/cuis.htm>

33. <www2.cr.nps.gov/e-rehab>

34. <www.cr.nps.gov/aad/nacd/>

35. <www.cast.uark.edu/local/online_map/>

36. <www.cr.nps.gov/csd/collections/parkprof.html>

37. <www.whitehouse.gov/WH/EOP/First_Lady/html/treasures/index3.html>

Long-Term Management and Accessibility of Archeological Research Data

PETER MCCARTNEY

Two decades of advances in computing have empowered the sciences to acquire, analyze and share vast quantities of information in digital form. Without standardization of protocols, documentation and commitments to maintaining data integrity, the value of our current holdings of electronic data is likely to depreciate rapidly. These issues place new responsibilities on archeological repositories to adapt to a world where the published literature is no longer our sole recording medium. This paper summaries procedures taken by the Archaeological Research Institute to address long-term data viability and to maximize the potential for networks to manage and disseminate research data.

The Archaeological Research Institute was created at Arizona State University in 1996 through an endowment from the Bureau of Reclamation to archive collections and associated data from four archeological projects in central Arizona. ARI has become a center for archeological data management, GIS research and Internet application development. In 1997 ARI assumed data management responsibility for the Central Arizona-Phoenix Long-Term Ecological Project, run by the ASU Center for Environmental Studies. The LTER program shares with ARI an emphasis on long-term data preservation and access, and has fostered an exciting mix of archeological and environmental research activities.

Data Archives at ARI

Since its beginning, ARI has archived and published three major archeological datasets from the Roosevelt Lake and the Lower Verde archeology projects.[1] Efforts are now in progress to include several more (Table 1). Most notable of these is a National Science Foundation grant to archive and document data files from the Teotihuacan Mapping Project. Procedures developed for archiving tabular data involve — 1) Acquire original data files in whatever format they are available. 2) Import the data files to an SQL database, checking for referential integrity and correcting errors where possible. 3) Reverse-engineer the data structure to produce a schema from which metadata are generated and published on the Web site, along with references and excerpts from any associated reports. 4) Export tables to ARI's SQL database server. Data are maintained as close to their original structure as possible, with only minor normalization as needed to more efficiently store the tables. Variables are not recoded to fit a master database structure as this would result in an unnecessary loss of original information.

ARI databases are accessible currently via an

Dataset	Creator	Status
Roosevelt Community Development Study	Desert Archaeology	Online
Roosevelt Rural Sites Study	Statistical Research	Online
Roosevelt Platform Mound Study	Arizona State University	Online
Shoofly Project	Arizona State University	In progress
Lower Verde River Project	Statistical Research	In progress
South Germany Survey Project	UCSB, ASU, Oberlin	In progress
Phoenix Indian School Project	Arizona State University	In progress
Teotihuacan Mapping Project	Arizona State University	In progress

Table 1. Datasets published at the Archaeological Research Institute.

HTML form interface that passes SQL queries to the database server, returning the results as either an html table or ASCII file. Metadata are available online as a guide to formulating queries. The interface provides an easy-to-use form for selecting data that hides much of the complicated table joining.

This year ARI has experimented with a GIS-based interface to these query systems to provide a more intuitive navigational guide to the Roosevelt Platform Mound Study database. Users can interactively pan, zoom and select sites and features to query, gaining a more visual appreciation for their spatial context at the same time. Through its partnership with the LTER project, ARI maintains a data library of GIS framework data on cultural, environmental and infrastructural coverages derived from sources such as Arizona State Land Department, USGS, Maricopa Associated Governments and Arizona Department of Environmental Quality. ARI has developed a searchable online SQL-based data catalog containing metadata based on ecology and geospatial standards. Metadata descriptions for archeological datasets now are being entered into this system to replace the static html-based metadata now in place on the ARI server.

ARI has invested some effort into digital archival strategies for non-tabular information. Photographic images have been digitized selectively and added to the database, although the primary application of this effort is for viewing (rather than archival) purposes. CAD and GIS data are maintained in native formats and can be accessed from the ARI lab computers. To facilitate Internet access, some data files are also available in pdf or in compressed zip formats. ARI has invested some resources in experimental technology that captures three-dimensional morphology and surface patterning of objects using a laser-scanner paired with a separate color scanner head. This produces a three-dimensional model of an object that can be stored, viewed, shared across networks and analyzed. A proposal recently funded by the National Science Foundation Knowledge and Distributed Intelligence program seeks to develop applications that will aid lithic refitting studies and help redefine parameters used in ceramic morphology studies.

ARI has formed a partnership with the Council for the Preservation of Anthropological Records to develop an online catalog to anthropological collections.[2] CoPAR has developed a metadata standard for describing the content, location and availability of anthropological collections. This system will provide a search tool for locating datasets within an international network of participating repositories.

A similar project at ARI is the AZSITE Cultural Resources Inventory, a collaboration with the Arizona State Museum, the Arizona State Historic Preservation Office and the Museum of Northern Arizona. AZSITE is a spatially-enabled client-server database of archeological and historic resources in Arizona that uses SDE, a server product from ESRI, to manage GIS data inside an MS SQL Server database, allowing large numbers of remote users in government, consulting and research institutions to access a database of 100,000+ sites and properties.[3]

ARI recently has turned its attention to the "gray literature" problem by partnering with CAZMAL, a consortium of Arizona museums and libraries, to develop its existing Web-based bibliography database into a shared, statewide resource. This project is designed to create a system that integrates with many existing research, public service and educational resources on the Web.

Infrastructural Needs

The directions taken by recent funding initiatives make it clear that if archeologists are to obtain sustained support for information technology, strategies must be developed within a broad model of future information systems. To think about future needs, we must consider current limitations.

Recovery and long term management of primary data. Among the most pressing problems facing all scientific disciplines today is the long-term survival of research data. The advent of electronic data management, although providing limitless capacities for storing and manipulating information, has initiated the growth of a tremendous body of primary data that exists exclusively in what is paradoxically the most imperiled medium of all-digital format.

Location and usability of relevant information. Preservation efforts are often directed at the basic task of recovering and stabilizing information in some lasting form. However, to really achieve the underlying goals of this effort, it must be possible to locate and make effective use of this information. Few studies attempt to use the limited extant archives of data because of the lack of documentation to guide their use. Vital information about the collection and processing of field data is often buried in unpublished records and "gray literature".

Diverse and dynamic state of information storage formats. Primary data sources differ in data storage schemas, file formats, communication protocols, etc., requiring further development of enabling technologies that permit query and retrieval from diverse systems in remote locations.

Data comparability. Archeological data are highly contextual, with the result that units of measurement, classification taxonomies and sampling designs often are not directly comparable.

Scaling observations to research at different spatial or temporal domains. Observations typically are made at the lowest scale of measurement, such as the feature, and must be synthesized and interpreted to express variability at higher scales, such as site or region.

Sharing data and applications within multi-institutional research collaborations. As research questions scale beyond the local region, the network of collaborators often grows beyond a single institution. Traditional media for data storage do not lend themselves well to remote collaboration. Even many electronic formats do not support use within a networked or multi-user environment.

Managing complex forms of information. Simulation modeling, GIS representations of settlement networks, CAD models of architectural features and three-dimensional images of artifact morphology all present new, complex forms of primary data to be managed.

An Information Management Plan for the 21st Century

Western science has depended for centuries on the published literature as its primary archiving medium. Responding to these challenges calls for the application of information technology beyond simple data storage and electronic publication to develop an active, globally integrated information network with the capacity to discover, access, interpret and process data fluidly across comparability and scaling barriers. Creating this infrastructure requires investing effort and resources into three broad areas that we might model as three layers within a continuum of increasing abstraction from data to information to knowledge (Figure 1).

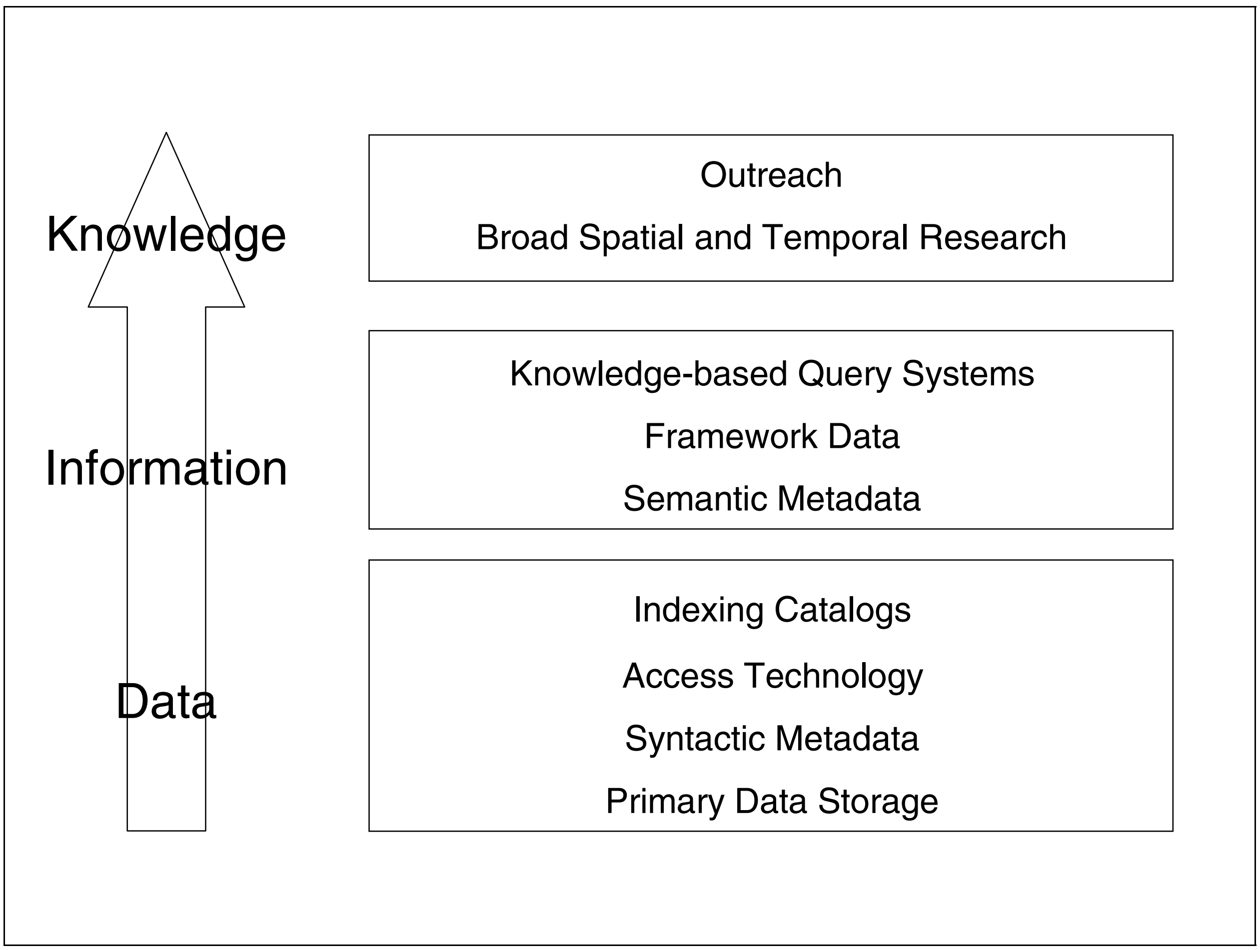


Figure 1. Model of a tiered information management infrastructure.

Data: Establishing Data Storage Network

The diverse range of information activities we see today will contribute to a global infrastructure for managing and accessing research data. This will require investments in several areas: data management strategies, metadata standards, access technologies and solutions for preserving digital data.

Data repositories:

Availability of data can be threatened by both short-term factors such as power or device failure or long-term factors such as media decay or format obsolescence. Solutions for the former involve regular backups to removable media, use of redundant subsystems, etc. Long-term availability depends on maintaining data in an online system with a plan for migration to new hardware and software. To successfully maintain data, an institution must have the resources and the commitment to manage and upgrade equipment and software and to maintain connectivity. While installing a data server and a Web connection is relatively easy and inexpensive, maintaining that connection over several years — let alone in perpetuity — is quite another matter. What is needed is a network of data repositories that can reasonably make this commitment.

Syntactic (Data-bound) Metadata

The term metadata refers to data that describe data. Metadata represent the key element to transforming archived datasets into useable research resources. In this level of the model, information about the syntax of the data — information that describes each specific dataset — is considered. This information is inextricably bound to the dataset and is thus expected to be stored and managed in close conjunction with the actual data.

In a seminal paper about the survival of ecological data, Michener and others[4] identify five levels of metadata description required to fully document an ecological dataset. These range from information about the research project that produced the data — names of investigators, sampling strategies, collection methods, etc. — to detailed attributes of the columns, datatypes and file formats of the data tables that were archived. A project at the University of Kansas funded by the National Science Foundation is developing a metadata standard for indexing and describing museum collections data; similar efforts to index electronic resources on the World Wide Web have been made by the Dublin Core initiative. [5]

There is a need for developing widely accepted standards for ecological metadata that go beyond this simple beginning. These metadata need to be developed following a modular approach similar to other well know standards efforts such as the W3 Consortium, FGDC and Dublin Core. Discrete working groups focusing on specific content domains would contribute towards a comprehensive standard that serves not to dictate research methods, but rather to effectively document the structure and design behind one's methods and observations.

Access technology

Our current data publication solutions, while effective, are inherently proprietary to the specific data content, storage and delivery system and thus time-consuming to develop. A layer of open access technology needs to be draped over this network of data repositories to facilitate the most fundamental search and query operations from a single agent using a single protocol. Several technologies currently in development point to the kind of solutions that will be included in this access layer. EXtensible Markup Language (XML) provides a language for creating machine-readable metadata. Client server search tools such as Z39.50 used by libraries and museums provide a platform-independent query language capable of searching multiple data sources on diverse hardware and software implementations. Newly proposed technologies such as the storage request brokering software under development at San Diego Supercomputing Center and the National Center for Ecological Analysis and Synthesis are designed to hide the platform differences among different storage systems behind a single universal interface that can be accessed by query applications.

Sustainability

If a data network system is to be sustainable, it is necessary to develop incentives and pathways for bringing data into the network at minimal cost. Despite metadata's vital importance, few active research projects take the effort to produce metadata for their research data. The cost of generating metadata often is prohibitive; and adequate guidelines for generating metadata do not exist.

Practices observed in other disciplines suggest several directions to pursue. One is to encourage funding and permitting agencies to endorse the submission of research data into knowledge repositories and to adopt a set of standards for this process. Another is to work with professionals to develop programs that create reward structures for data archiving and documentation. One such program, created through cooperation between the Ecological Society of America and the Oak Ridge National Laboratory, developed a peer-review process for datasets and associated metadata, with successful submissions receiving honorable mention in the ESA journals. Introduction of the concept of peer-review for datasets is consistent with recent changes to the National Science Foundation grant guidelines that request listings of all data from prior NSF research that have been published online.

Finally, the cost and difficulty of creating metadata might be mitigated by developing freely distributable tools that automate the documentation process through reverse-engineering of data files and use of "wizard" forms that query the investigator for information similar to the way tax wizards gather financial backgrounds. Through cooperation with archeologists from Desert Archaeology, Inc. and the Arizona Department of Transportation, ARI is working to develop a simple remote entry tool for preparing archeological project metadata for submission to ARI's data archive. With sponsorship from regulatory agencies and major funding institutions, completion of such standard metadata eventually should become a routine requirement for all permitted research.

Indexing Catalogs

As the corpus of online data resources grows, the need for efficient indexing and searching far outstrips the capacity of static and unsophisticated aids such as html link pages and Webcrawler-based search engines. Current efforts by CoPAR, the Dublin Core and the Encoded Archival Description projects to build structured metadata indexes are building a valuable infrastructure for navigating the growing network of digital data. A model for indexing efforts such as CoPAR is the National Spatial Data Infrastructures network of clearinghouses for geospatial datasets.

Information: Integration and synthesis of data

The interface between the data storage systems described above and the kinds of synthetic research questions we wish to accommodate within our broad infrastructure is largely undeveloped. Several key components are likely to see significant attention over the next decade.

Semantic (Query-bound) Metadata

One component is a set of standards for decomposing research questions into smaller elements that can be documented in standardized, machine-parsable form. We can refer to these elements as semantic metadata because they concern themselves not with the organization of information but with its meaning. Semantic metadata are query-bound in that they provide a means of documenting our diverse units of inquiry, just as the more familiar syntactic metadata documents are diverse units of observation. They document the calibration curves, classification schemas or processing steps necessary to transform and scale data from different sources to address a general question.

Develop framework data to summarize regional patterns.

There is a growing need for regional summary — or framework — data as a means of making archeological results available and useable to non-archeologists and to archeologists needing synthesized data for textbooks, atlases, basemaps, etc. These data would be maintained through software that reprocess the data as new information is made available, thus improving the quality of data used in everyday decision-making. ARI is collaborating with several partners to produce online framework data resources. These projects, some still pending funding, include a database of petrographic analyses in Arizona, an inventory of rock art for northwestern Arizona and a comparative database of Mesoamerican figurines from Teotihuacan.

Develop tools and procedures for automated integration of data

Few tools exist today to facilitate the task of synthesizing data from diverse primary sources. It is reasonable to begin thinking about the application of technologies such as knowledge-based systems that could receive input using some query language, access both semantic and syntactic metadata and perform a certain amount of query, evaluation and processing of primary data prior to returning a result. Very simple applications of expert systems have been developed to perform tasks such as classify lithic artifacts by their attributes. A recent National Science Foundation proposal submitted by ASU hopes to produce even more robust applications.

The key to developing such tools lies in two areas. The first involves extending our partnerships with expertise in sophisticated computer technology such as artificial intelligence, expert systems and neural networks. These are not strengths traditionally found within archeology nor are the funding sources for these disciplines familiar ground for us. The other area involves developing a language for encoding both syntactic and semantic metadata in machine-readable form such as Resource Description Format, an XML-based standard gaining wide industry support. Current metadata implementations rely heavily on open text representations for information such as classificatory systems, measurement parameters, analytic procedures, etc., and are not machine-readable.

Knowledge: Promote and support research collaborations integrating information at broad spatial or temporal scales

Funding strategies for information technology during the last two decade have focused on building infrastructure with the notion that "if you build it, they will come." The new crop of initiatives shows much more concern with ensuring that our data products are of significant value to current and future research — that is, new proposals are expected to provide *application*, not just availability, of data. We need to challenge current modeling and synthetic research to make more extensive use of archived primary data so that we can create both incentives and guidelines for developing the infrastructural components I have outlined above.

Archeology needs to take a pro-active role in establishing the long-term value of its data products. While we generally recognize our own research goals as the guiding forces behind our work, we need to recognize that general scientists, public policy makers, businesses, the legal profession, K-12 educators and even the entertainment industry, all make use of information about the past and that there can be serious implications if those users make uninformed decisions based on faulty, outdated or incomplete information. The partnership between ARI and the Long Term Ecological Research project was born, in part, out of a desire to see data and research methods developed in archeology applied in new interdisciplinary research contexts.

Finally, information managers in archeology need to follow the example of their counterparts in ecology and earth sciences in recognizing that the information systems of the future will depend on standards resulting from collaborative effort. Individual solutions alone will not carry us beyond the lowest level of integration outlined here. Mechanisms such as the LTER inter-site data management committee, the PACRAT working group of the Mohave Desert Ecosystem Initiative or the proposed Bio-Informatics Consortium described in an up-coming LTER white paper[6] all serve to facilitate cross-fertilization of ideas and the development of standards for data communication.

Conclusions

We are poised at the beginning of what promises to be an exciting time for information management in the social and environmental sciences. Bandwidth is getting wider, IT funding initiatives are getting bigger and computers keep getting cheaper. The only thing standing in our way is our own ability to think far enough ahead to build the kind of infrastructure that will let us apply data to research and to real-world problems in the manner we want to be able to.

Acknowledgments

While I have chosen the particular words to express the issues raised in this paper, I must acknowledge that many of the ideas have been born from recent discussions with members of the Department of Anthropology, Arizona State University, the Long-Term Ecological Research Data Task committee, the National Center for Ecological Analysis and Synthesis and the San Diego Super Computing Center.

Notes

1. *Roosevelt Rural Sites Study Research Database*. Archaeological Research Institute. <archaeology.asu.edu/data/rss>. 1997.
Roosevelt Community Development Study Research Database. Archaeological Research Institute. <archaeology.asu.edu/data/rcd>. 1997.
Roosevelt Platform Mound Study Research Database. Archaeological Research Institute. <archaeology.asu.edu/data/rpms/>. 1998.

2. *CoPAR Guide to Anthropological Records.* <archaeology.asu.edu/cgar>.

3. <archaeology.asu.edu/azsite>.

4. W.K. Michener et al, “Non-geospatial Metadata for the Ecological Sciences,” *Ecological Applications* 7 1997: 330-342.

5. <purl.org/dc>.

6. McCartney et al, “The Future of Bioinformatics in LTER” (Draft white paper prepared for the LTER Network Office, 1999).

Coming to Terms with the Information Age in Archeology

MARK ALDENDERFER

My thesis in this paper is simple: although I believe we have made real progress since 1994 in our ability to disseminate data effectively, we as a discipline have yet to give serious attention to infrastructural concerns that must be addressed as we seek to integrate more fully electronic means of data distribution into archeological practice. I shall discuss five issues: 1) the reality of funding, 2) the question of audience and ends, 3) the tyranny of standards, 4) the necessity of continuity, and 5) the innovations that will matter most.

As recently as 1994, it was possible to ask the question "Has archaeology remained aloof from the information age?" in all seriousness.[1] The answer offered by the author was an affirmative and he argued that although a number of innovative uses of computing had appeared in archeology, most of those responsible for conducting archeological research and ultimately disseminating archeological data preferred to publish in "the conventional manner" — that is, on paper in articles, books and monographs. Perhaps of greater interest was his belief that it was becoming increasingly necessary for those who receive archeological data for storage, such as museums, record centers and libraries, to become innovators and develop new and exciting ways to deliver archeological information to the public. One of his main points was that these institutions, often with little participation by those involved in data creation, had to grapple with the emergence and definitions of standards on how to describe data, the evolution of new and different modes of publication and ways to make information accessible to end users. Although his definition of the "public" is quite limited, his opinions nevertheless are a good departure point for examining just how far the discipline has come in the past five years in delivering archeological information to its various constituencies.

The Reality of Funding

Everyone knows that computing power, software, storage and bandwidth have become much cheaper over the past decade. We still have to buy all of it, of course, and often we find ourselves being pushed to keep up with extremely rapid changes in information technology that threaten to make our existing investment obsolete. An often ignored or overlooked cost of computing, however, is that associated with the deployment of a system of IT support that keeps all of the component parts organized and in good working order. Staff generally translates to "people," and as numerous commentators have observed, it is increasingly difficult to find and retain IT staff.[2] This problem shows no signs of abating. Other costs of IT support include those related to data acquisition and evaluation, entry and metadata construction.[3]

These costs of computing create a real dilemma in archeology, because by any measure we are not a well-funded discipline. Our available resources are generally spent on basic research, analysis, curation and heritage management. Despite a growing recognition that possession of an effective IT strategy is required for individuals as well as organizations, finding resources to implement the strategy is very difficult. Childs (this volume), describing Federal archeology on the Internet, makes a telling point: "The bottom line is that most Federal agencies with archeology programs do not have a formal Web design and development infrastructure for cultural resources, let alone archeology…Regular maintenance is difficult and new product development without extra help and money is even more difficult."

This scenario is replayed in academic contexts. At my university, I have observed that while it is only moderately difficult to get new equipment, it is almost always close to impossible to get new IT staff. This is at least partially reinforced by our funding agencies. In the Social, Behavioral and Economic Research directorate of the National Science Foundation, which hosts archeology, there are a number of competitions that support the purchase of new equipment for both research and educational ends. But there are no competitions that finance the addition of staff aside from those that support laboratories, a topic I will return to below. While it is easy to make the case that NSF is not in the business of supporting infrastructure, this acknowledgment does not make it any easier to achieve our goals.

The reality of funding, then, is this: if digital dissemination of archeological information is to become a reality, the true and full costs of supporting an IT strategy for individuals and institutions must be recognized. At a minimum, such a strategy must involve regular infrastructure — hardware, software, bandwidth, etc. — and upgrading as well as maintaining an adequate level of support staffing.

However, if we are already stretching our limits in terms of current levels of support, how will this be done? We either find new resources or we reallocate existing resources to meet these new priorities. Although we could maintain the status quo — get done what can get done given what we have, rely on the goodwill of volunteers, etc. — so doing will mean that we will fall even further behind what might be useful and desirable to achieve our ends. And while I am always in favor of lobbying for new resources, it may be more realistic and practical to reassess our priorities. To do this, we must then consider our audience and ends.

The Question of Audience and Ends

The question "What are we disseminating and for whom?" should dominate our consideration of priorities. We all know that a data archive is developed for different ends than is a Web site created for a classroom. But one of the seductions of digital publishing is the belief that with just a little more effort, material presented to one audience or forum can be transformed into material for another. While it may be the case that from a purely technical perspective information can be reformatted and made available in a different package, this does not mean that the repackaged information is necessarily **useful** to that new audience. The analogy to print publication is clear: an author might choose to prepare a scholarly monograph for his colleagues, and it is thus written in a language and presented in a manner that reflects the expectations of that audience. However, it requires a separate intellectual act to write a book suitable for the educated lay public. Why? Because the public's expectations for content, style and presentation are very different from those of specialists; a mere repackaging of the monograph is unlikely to achieve the desired result. And while I grant that things difficult to do today will become tomorrow's routine task, the desire to do more will always be a tension since technology also will continue to offer ever greater possibilities.

Another growing conviction is that since technology offers us storage capacity — or virtual reality, complex three-dimensional renderings, high capacity bandwidth or other IT marvels — undreamed of even ten years ago, we should strive to use it in every instance. Recognition of this temptation arose, for instance, at a recent workshop on digital publishing sponsored by the Digital Archaeology Laboratory at the University of California at Los Angeles' Institute of Archaeology, during a discussion about the ideal content of a digital monograph. Some argued that a digital monograph could contain "all of the data" generated by a project and that it would ultimately represent a kind of archive offering both interpretation and data presentation. Others present were less sanguine, and suggested that such a monograph would serve neither end particularly well. If one wants an archive, there are better models for it than a monograph format. Although all present agreed that a digital monograph could be anything its authors wanted it to be, a general consensus emerged that defined a digital monograph as a traditional, but value-added product, the production of which was contingent upon the demands and expectations of the audience. Here "value-added" refers to the extras, like sound, color image presentation and innovations in the mode of navigation, that technology allows us to add to our books. But the point remains: our products — Web sites, archives or monographs — must be content, not technology, driven.

This brings us back to the reality of funding. Audiences for many types of archeological information are very small, in some cases no more than a handful of specialists worldwide. Complex data management projects in these instances, such as making digital large quantities of field notes and records from older projects, are very expensive and can thus be viewed as terribly wasteful of resources since so few benefit. Does this mean, then, that only those projects likely to reach large audiences, such the Ancient Architects of the Mississippi Web site,[4] should be digital? Surely, the answer must be "no" since we have an ethical obligation as archeologists to publish and preserve our data. But *must* we preserve it digitally? Strong arguments can be mustered for both pro and con. However, given current priorities that emphasize basic field research as opposed to collections-based projects, digitizing very large collections of original notes, records and other data that only appeal to small audiences will be difficult to justify.

The Tyranny of Standards

The computing world is a world of standards. Although I imagine that it would not be impossible to chart all of the standards that govern IT, it would be a Herculean task. Standards make computing possible because they facilitate communication — without standards, there would be no communication. Standards govern everything — from how words, sounds and images are translated into electrons, how those electrons are routed across networks, how the electrons are stored by media as disparate as floppy disks through flashcards through bubble memory and how the electrons are printed onto paper or other media. Standards govern how to describe data, and how to link data together so that other sets of standards recognize them. Standards, then, govern every aspect of IT from the most technical domain of hardware design and implementation to the broadest level of content definition and identification.

National and international bodies of all kinds — not to mention trade associations; business, academic, and interest-group consortia; and even individual companies — all define and then attempt to codify standards. Some standards have the imprimatur of government and are widely followed. Other standards are created by treaty agreements between nation-states as a part of their trade negotiations. Still other standards are less formal, but no less powerful. Consider the near-monopoly of the Windows operating system for desktop computers. No organization has forced the user to buy Microsoft, but instead, computing in North America has evolved to fix Windows as a *de facto* standard. Similar *de facto* standards of the recent past have been the 3.5-inch floppy disk (now being replaced by a number of rival candidates), CD (soon to be displaced by DVD) and the QWERTY keyboard (still going strong).

The tyranny of standards has two facets: 1) you do not control them, yet you must work with them as they dictate, and 2) they will change, and you will have to change with them. Both of these conditions lead back to the inescapable reality of funding: it is expensive to keep up with standards, but you have no choice if you expect to communicate. This is not a question of buying the "latest, fastest, greatest CPU" or keeping up with the Joneses, but a simple matter of keeping abreast of change and coping with it. Consider this example from the world of metadata standards:

"After conducting an analysis of the scope of the project, it was determined that database modeling should focus on the specific metadata requirements necessary to support the Federal Geographic Data Committee's (FGDC) Content Standards for Digital Geospatial Metadata (CSDGM), the Directory Interchange Format (DIF), the Government Information Locator Service (GILS), and the Earth Observing System Data and Information System (EOSDIS) Information Management System (IMS) data order requirements. As a secondary exercise, the data model will be extended with some additional content elements from the Anglo-American Cataloging Rules (AACR)/Machine-Readable Cataloging (MARC), and both MARC and Standard Generalized Markup Language (SGML) will be incorporated as data element tags into the data model."[5]

This statement was issued by CIESIN, the Consortium of International Earth Sciences Information Network, a not-for-profit, non-governmental organization established in 1989 at Columbia University to "help scientists, decision-makers, and the public better understand their changing world. CIESIN specializes in global and regional network development, science data management, decision support, and training."[6] Note that CIESIN itself does not create the standards for metadata per se. Instead, they coordinate projects that are required to use the standards set by others in order to comply with the terms of contracts. In the quoted example, CIESIN is describing the construction of something called the "Unified Metadatabase," used to describe the spatial features of demographic, health status and other data, and which in this case uses standards set by three US Federal government programs (FGDC, GILS, and the EOSDIS IMS), a *de facto* data exchange format standard that was created by NASA (DIF), a Library of Congress-supported library coding system that has broad-based English language and academic support (MARC) and a text markup language accepted by the International Standards Organization in 1986 (SGML).

I hope the point of this example has not been lost in the maze of acronyms. It takes time, resources and staff to keep up with ever-changing standards and investing resources to stay current with standards may well diminish resources destined for basic research. Changes in hardware, software and communications standards will require investment in infrastructure and the development of a planning process. We should not, however, ignore the emergence of metadata standards because, in the long run, knowledge of and compliance with these standards will structure **every** digital means of information distribution likely to appear in archeological practice. The question of how shall we seek guidance leads to the next issue, that of continuity.

The Necessity of Continuity

Until recently, the question of continuity has been relatively straightforward for archeologists concerned with their ethical obligations for data dissemination. The relevant repositories have been the library (which houses the monograph), the archive (which houses the unpublished records) and the museum (which houses the objects of study). Digital publishing, however, changes all of this. Where does your Web site go when it's time for it to retire? Do you keep a copy as a record? Should you? Who keeps your database online? For how long? And who will pay for the costs of keeping up with changes in standards? You? The organization? The government? No one?

Answers to these questions are emerging very slowly in archeology, but we can outline some of the directions our field is likely to take. Although costs will always constrain what we do, continuity in digital publishing depends more on what it is we are publishing and what kinds of obligations we have to preserve these products. Some things probably don't deserve preservation in perpetuity — classroom Web pages are a good example of ephemera. But we'd like to preserve most other digital products, like primary data records, monographs or other published works.

The preservation of a digital monograph is a good place to begin our discussion. As I noted above, such a monograph can be anything the author wishes and given existing technology, such publications can contain interpretations and all of the primary data the author sees fit to include. At present, the only two digital formats that could accommodate this enhanced vision of a monograph are either CD or DVD "books" and Web sites. I know that a book, if published with acid-free paper and kept under reasonable conditions, will last many decades. A CD, however, has a much shorter shelf life, perhaps only two decades at most. Even if CD readers are still available in three decades, will someone be able to "read" my CD book?

Changes in hardware and software standards may even affect content. If I have chosen to publish my data records in table format, it is likely that some program will exist to translate my data into a new format, although it is likely that the original formatting will be lost. But if I represent my data by means of some complex construction, such as a three-dimensional rendering of a site or building, there exists the real possibility that changing standards may well make it impossible to capture that rendering accurately. That part of my book, then, will be lost even if some entity exists to migrate my original work into a new form.

Web sites offer similar problems. We really have no idea what 'in perpetuity" means with regard to a Web site, but if modern Web practice is any guide, the frequency with which links to pages are returned as non-existent does not bode well for long-term maintenance of complex sites. Having greater storage capacity at the site and more bandwidth available to display more complex images or information more rapidly does not guarantee longevity.

How continuity is achieved is scale-dependent and closely related to the type of digital document produced. Individuals must consider their own IT strategies. You are ultimately responsible for migrating data to new formats and platforms if you choose to store them digitally. Likewise, it is your responsibility to keep your digital archives of papers online should you wish to do so. However, as we move away from the desktop, individuals begin to lose control of their products and it is here that existing institutions must be adapted to new ends or wholly new organizations developed to ensure continuity.

An existing institution undergoing significant reorganization to meet the digital challenge is journal publication. Publishers are experimenting with a plethora of organizational and business models to implement a digital dissemination scheme. The Institute of Electrical and Electronics Engineers, for example, has eschewed print publications altogether and has an ambitious timetable for migrating all of its journals to an online format by the start of the next decade. Traditional publishing houses have created hybrid models, where both print and digital copies of journals or in some cases individual articles are made available through different subscription formats. Not surprisingly, libraries have become concerned with how they maintain their place as repositories of journals as these changes take place. *Internet Archaeology* is the first major archeological journal in a wholly online format.[7] Sponsored by a consortium of British universities, the journal is presently free of charge and is distributed through servers maintained by the consortium. Subscriptions are planned, which may mean a change in the form of distribution as well. Regardless of what we publish, however, continuity requires that some organization will assume the costs associated with access to a product and will provide the resources for continuous migration of the products as standards change.

The maintenance of digital data archives merits special consideration. In this instance, I mean the term "data" to refer to primary records that describe field contexts, objects, images and other products of field and laboratory studies. As we all know, archeology is a destructive process, and we rely almost wholly upon the paper records, photographs and drawings of the specific field contexts from which the objects of our study have been taken. Without context, we have in a real sense no data. If we embark on a program of digital capture and permanent archiving of these data, we have to guarantee with whatever certainty we can muster that we will in fact maintain these records in perpetuity. This is certainly the implicit promise in a paper-based archive, although it is clear that because paper has longevity, it has been possible in great part to defer the upgrade of the paper products to new formats. But we know all too well that our digital products have a much shorter shelf life and the making of the promise has very tangible and near immediate impacts.

Digital archives, then, must have a very real organizational imprimatur that guarantees continuity. Otherwise, if we lose *these* data, they are lost forever. Many groups have been concerned with these problems and a few sources of useful information include Beagrie and Greenstein,[8] which contains a number of excellent case studies on how archiving institutions must develop effective IT strategies, Day,[9] which includes a comprehensive bibliography on digital archiving and Conservation OnLine,[10] an online catalogue of information for professionals in the digital archiving field that has much to offer anyone interested in the topic.

What kinds of institutions currently support digital archives? These range from international scientific organizations such as the International Council of Scientific Unions, which supports the World Data Centers — which is itself composed of national and academic bodies that now warehouse primarily geological, geophysical and similar data — governmental agencies of all kinds at all levels — from the local to the global — professional societies and avocational groups. Among the organizationally supported data archiving projects in archeology are the National Archaeological Data Base[11] — supported by the National Park Service — the Archaeological Data Archive Project — supported by the Center for the Study of Architecture at Bryn Mawr University[12] — and the Archaeology Data Service[13] — sponsored by a consortium of universities, museums and government programs in the United Kingdom.[14] Although these organizations are a good start, it is clear that archeology is far behind most disciplines in the way in which it has approached digital data archiving and dissemination, and this in turn leads to my final point, which is concerned with the significant innovations.

The Innovations That Will Matter Most

If archeology is to join the information age in a serious way, we must see innovation in three domains: data acquisition, education and perception.

Basic field recording techniques in archeology have changed very little over the past 100 years and involve the use of a combination of paper forms, notebooks, graph-paper drawings and standard 35mm and large format photography. While these techniques are reliable, they are very limiting, especially as one moves from the field to analysis into data publication, presentation and archiving. Field drawings must often be redrawn and digitized by hand for integration into advanced geographic information systems. These same field drawings must also be linked by hand to computerized databases that describe their contents. Handwritten field notes rarely are transcribed and searched electronically for information and forms, while they always contain important information, have to be summarized and described and their content re-transcribed into other paper or possibly digital records. Field notes and forms are searched visually by flipping through ring binders or file folders. Slides, prints and negatives can be integrated into databases, but it is difficult to integrate them easily into sets of field drawings and maps in a consistent manner. And while many archeologists have begun to digitize these data so that modern IT tools can be used to examine them in a more rapid manner, the costs of this post-hoc approach are very substantial and further, they tend to introduce new sources of error into these primary data. Indeed, many archeologists have come to believe that traditional field recording methods substantially slow the pace of analysis and certainly the publication and archiving of the results of field research.

With the advent of pen computers, digital theodolites, and other in-field data recording devices, many of the technical challenges of digital primary data recording are now at least surmountable and as these devices improve — as they inevitably do — the costs of so doing will diminish as well. This implies that in-field digital recording of primary archeological data can become standard archeological practice and while it may not replace paper in all circumstances, it will certainly become more widespread. However, this should only occur if students are trained effectively in these technologies and if the use of these technologies is seen as only part of a comprehensive IT strategy and not simply an end in itself.

As I pointed out in a recent review paper on quantitative methods in archeology,[15] we as a discipline need to focus more of our attention on how we train our students to use these methods. This argument applies with equal force and relevance to IT. Learning how to use a computer or specialized software is not the same as understanding how to integrate IT effectively into the research process. Although there are many students who learn how to use various types of IT on their own, I have argued that if we want to train true innovators, we will have to provide the appropriate educational context to do so. One such program is the M.Sc. in Archaeological Computing at the University of Southampton.[16] Topics covered include digital drawing and imaging, database systems and geographic information systems. Programs like this will have to become more numerous if we expect to have our "own" experts adapt IT to archeological ends. All of this, unfortunately, costs money, and from where will it come?

This, then, brings us to the final innovation that will matter most — a modification of our funding policies. Frankly, I don't expect that this will happen, but I think it is a message that needs to be delivered regardless of the reception. If digital dissemination of information is to become commonplace in archeology, we must build at every level — from the individual to the global — an effective IT strategy. This will cost money and unless we get new resources, we must create the political will to re-allocate existing resources. Individuals have their part to play in this, but as I have argued, if we are to have real continuity in data dissemination, we will have to look to the creation of organizations that will be charged with these responsibilities. Leadership is required at all levels and we can attempt to enlist the services of the Society for American Archaeology, the Archaeological Institute of America, the National Park Service and the National Science Foundation. The latter two may be the source of funds, but we will need the backing of our professional and academic societies to provide part of the push. NSF has already begun to fund digital data dissemination projects that target archeology, but we need to see more of these.

I also believe that the development of a new conservation ethic for existing data much as has developed in our field for the preservation of archeological sites themselves must be a necessary part of this strategy. Indeed, the archiving of all data records — whether digital or not — is taking on new importance as many professional and academic societies of archeologists, such as SAA, AIA and the Register of Professional Archeologists, among others, promulgate strict archiving standards from an ethical perspective.[17] The demands of digital preservation, however, create even more urgency for such standards to be created. Once again, we can begin with individuals who take their ethical responsibilities seriously and will either develop their own IT strategies or will deposit their data with existing organizations. This will require a change in mindset, obviously, as well as the reward structure in both academic and professional arenas and we should begin to get equal credit for preserving, as well as publishing, the results of our labors.

Back to Booth. We've not remained aloof from the information age, but we still have far to go to recognize what the full implications of going digital will be to our field. I am certain that we have a digital future, but what form it will take is still not clear to me. The optimist in me sees NSF-funded national centers for digital data preservation; the pessimist sees the Dead Media Project page on the Web.[18] You choose.

Notes

1. B. Booth, "Has archaeology remained aloof from the Information Age?," In *Computer Applications and Quantitative Methods in Archaeology 1994*, eds. J. Hugget and N. Ryan (Oxford: British Archaeological Reports, International Series 600, 1995), 1-12.

2. "Current Issues for Higher Education Information Resources Management." *CAUSE/EFFECT* 20.4 (1998): 4-7, 62-63. <www.educause.edu/ir/library/html/cem9742.html>.

3. V. Canouts and S. T. Childs, "National Databases: Promise and Postscript" (Paper presented at the Society for American Archaeology Annual Meeting, Seattle, 1998).
A. Wise and P. Miller, "Why Metadata Matters in Archaeology," *Internet Archaeology* 2 (1997). <intarch.ac.uk/journal/issue2/wise_index.html>.

4. <www.cr.nps.gov/aad/feature/>.

5. *CIESIN Metadata Guidelines* <www.ciesin.org/metadata/TOC/init.html>. 1998.

6. Ibid.

7. M., J. Heyworth, A. Vince Richards and S. Garside-Neville, "Internet Archaeology: A quality electronic journal," *Antiquity* 71 (1997): 1039-42.

8. N. Beagrie and D. Greenstein. *A Strategic Policy Framework for Creating and Preserving Digital Collections*. Arts and Humanities Data Service. 1998. <ahds.ac.uk/manage/framework.htm>.

9. M. Day. *Preservation of electronic information: a bibliography*. 1997. <www.ukoln.ac.uk/~lismd/preservation.html>.

10. <palimpsest.stanford.edu>.

11. <www.cr.nps.gov/aad/nadb.htm>.

12. H. Eiteljorg, "Electronic Archives," *Antiquity* 71 (1997): 1045-57. <csaws.brynmawr.edu:443/web1/adaparchive.html>.

13. <ads.ahds.ac.uk>.

14. J. Richards, "Preservation and Re-use of Digital Data: The role of the Archaeology Data Service," *Antiquity* 71 (1997): 1057-59.

15. M. Aldenderfer, "Quantitative Methods in Archaeology: A Review of Recent Trends and Developments," *Journal of Archaeological Research.* 6.2 (1998): 91-120.

16. <www.arch.soton.ac.uk/Prospectus/ Computing/>

17. N. Parezo, and D. Fowler, "Archaeological Records Preservation: An Ethical Obligation," in *Ethics in American Archaeology*, eds. M. Lynott and A. Wylie (Washington, D.C: .Society for American Archaeology, 1995): 50-55.

18. <www.islandnet.com/~ianc/dm/dm.html>.

Contributors

Mark Aldenderfer is professor of anthropology at the University of California-Santa Barbara. He is interested in the ways in which humans have used high mountains and plateaus in the past and has conducted fieldwork in the three major highland regions of the world — Ethiopia, Peru and Tibet. He supports his field interests with a number of methodological specialties — including the use of GIS in both regional and intrasite research, archeological statistics and lithic analysis. He has become increasingly concerned with developing novel technical approaches to the acquisition of primary field data in digital form as well as ways of publishing, storing and archiving digital data.

David L. Carlson is an associate professor of anthropology and the department head at Texas A&M University. His research involves quantitative applications to archeological sites ranging from the early Holocene to the nineteenth century in North America and Mesoamerica. On the Web he is the listowner of ARCH-L — an electronic conference for archeologists; the author of "Frequently Asked Questions About a Career in Archaeology" and he maintains "Anthropology in the News" — a Web site that provides links to current news stories that relate to anthropology.

Mary S. Carroll is information management coordinator at the National Center for Preservation Technology and Training — an office of the National Park Service, Department of the Interior — where she manages the Information Management and Publications Support project types of the Preservation Technology and Training Grants program. She received her MA in anthropology with an emphasis in archeology from Arizona State University. Her interests include the analysis and design of computerized database management systems, the use of Internet technologies to disseminate preservation-related data and the long-term preservation of digital data.

S. Terry Childs is an archeologist for the Archeology and Ethnography Program and leader of the Cultural Resources Web Team for the National Park Service. She received her BA in Anthropology from Brown University and an MA and Ph.D. in Anthropology from Boston University. Her research over several decades has focused on the anthropology of technology, primarily in Africa. In recent years she has also turned her attention to education on the Internet and archeological curation. She is a research collaborator at the Smithsonian Institution and a research associate at the R. S. Peabody Museum in Andover, MA.

Christopher Chippindale is a curator in the Cambridge University Museum of Archaeology & Anthropology, England. His own field research largely concerns rock-art in Europe and in Australia. In his other trade as editor, he edited the international journal *Antiquity* (1987-1997) and now chairs the Publications Committee for the Society for American Archaeology. He is interested in how the changing technology of communication itself shapes the form that knowledge takes.

Harrison Eiteljorg, II, is a classical archeologist specializing in the architecture of ancient Greece. He has been active in exploring computer applications in archeology, founding the Center for the Study of Architecture (CSA) — which he directs — to encourage the use of computer-aided design software in archeology and architectural history. He is also the director of the Archaeological Data Archive Project (ADAP), an archive of digital data generated in the process of archeological research.

James A. Farley is chief technology officer for the University of Arkansas J. William Fulbright College of Arts and Sciences and the technical director at the University's Center for Advanced Spatial Technology. He is research associate professor in anthropology with faculty affiliations in Environmental Dynamics and Middle Eastern Studies. Farley is a co-founder of the Open GIS Consortium and has conducted a series of international lectures on GIS, technology, interoperability and business-technology integration over the past five years. He is a 1999 Computerworld-Smithsonian Laureate for his research in spatial data warehousing. Farley is currently active in the international disaster management community and the Global Disaster Information Network. He is working to establish a global network of technology centers that support faculty and student research directed towards broad-based environmental issues.

Hugh Jarvis is a graduate student of the University at Buffalo. He is completing an MLS in the School of Library and Information Studies and a Ph.D. in Anthropology. His interests include lithic sourcing, prehistoric hunter-gatherer territoriality and electronic publishing. He has established a variety of Internet and Web-based resources, including the Anthropology

Review Database, the Worldwide Email Directory of Anthropologists and the "lithics site." He runs e-mail discussion lists Anthro-L, Lithics-L and ArchComp-L.

Bart Marable is the creative director and principal of Terra Incognita, an interactive design firm that specializes in producing educational and entertainment Web sites. He has produced projects for the National Park Service, the National Geographic Society, the Smithsonian Institution, the MCI WorldCom Foundation, NASA, and Royal Caribbean, among others. After completing a master's degree in modern European history at Louisiana State University, Bart formed Terra Incognita in order to continue his work in developing educational interactive projects. Since the opening of the firm in May of 1995, his work at Terra Incognita has been recognized by awards from *Communication Arts, How, Critique, Step-by-Step*, *High Five*, *Project Cool*, and IPPA. His work has also appeared in the *New York Times*, *National Geographic Magazine* and *USA Today*, as well as on CNN and National Public Radio. He is also a frequent contributor to *Web Techniques* on visual design issues.

Peter McCartney is assistant research professor in the Center for Environmental Studies at Arizona State University. He serves as Information Manager for the Central Arizona - Phoenix Long-term Ecological Research project and as director of a three-year NSF-funded project to develop an information infrastructure for environmental data resources in Arizona. He also is a PI or co-PI on several other anthropological database projects including archiving data from the Teotihuacan Mapping Project, the AZSITE Cultural Resources Inventory and the Council for the Preservation of Anthropological Records. Prior to joining the Center, McCartney was information manager for the Archaeological Research Institute at ASU where he developed the Internet database publications described in this publication.

Donald H. Sanders is trained and educated as an architect, architectural historian and archeologist. His special interest is the application of nontraditional methods — including advanced computer graphics, virtual reality and behavioral science techniques — to the study and presentation of architecture of the past, pushing the boundaries of conventional archeological interpretation. Professional publications and conference papers have covered such topics as the historiography of the study of architecture by historians and archeologists; alternative approaches — including those from semiotics, environment-behavior studies, ethnoarcheology, and human geography — to the study of architecture in archeological contexts; and the use of computer-aided techniques for the collection, analysis and dissemination of information about ancient architecture. He has taught college courses in archeology and architectural history and was architecture editor for the J. Paul Getty Trust *Art and Architecture Thesaurus*. For the past four years he has been president of Learning Sites, Inc., which designs and develops educational and research software using interactive three-dimensional digital models that are based on actual archeological evidence.